VOLUME XXX OCTOBER, 1939 PART 2

THE FACTORIAL ANALYSIS OF ABILITY

I. THE PRESENT POSITION AND THE PROBLEMS CONFRONTING US[1]

BY GODFREY H. THOMSON

Moray House, University of Edinburgh

I. *Why do psychologists want factors?* (pp. 71–72).
II. *The problem of factorial analysis* (pp. 72–73).
III. *Reproducing the variance* (p. 73).
IV. *Reproducing the correlations* (pp. 73–75).
(1) *Minimal rank methods* (pp. 73–74).
(2) *The Spearman school* (p. 74).
(3) *The Thurstone school* (pp. 74–75).
V. *Reciprocity of tests and persons* (pp. 75–76).
VI. *Various other matters* (pp. 76–77).
References (p. 77).

I. WHY DO PSYCHOLOGISTS WANT FACTORS?

I PROPOSE to set out what appear to be the main principles (some of them incompatible) of the several systems of factorial analysis which co-exist at the present day. First, however, let us ask why psychologists want to have factors at all. The following reasons are given:

(*a*) If factors can be so chosen that a few of them give a good approximation to the information given by the large number of tests, there is obvious economy in their use.

(*b*) Orthogonal factors, that is, uncorrelated traits, have the advantage for scientific thought that they are independent; and they lead to simpler formulae. It should be remarked, however, that none of the human traits naturally named by naive man are uncorrelated, and that he is usually unable to realize the independence of the factors offered him by the psychologist, unable to realize for example that a man of high v (the verbal factor) is just as likely to be a man of low as of high g.

(*c*) There is a feeling that factors may be more enduring entities than the innumerable and changing tests used to find them. They come to be

[1] A contribution to the symposium presented at the Extended General Meeting of the British Psychological Society held at Reading in April 1939.

looked upon as the things in terms of which tests are described, although really of course it is the factors which are described in terms of tests.

(*d*) It is an easy transition to look upon the factors as actual and real. It is of the nature of man to deify or reify forces and powers behind phenomena, and we are all subject to this urge, which is, I think, a large part of the explanation of why factors are so acceptable to so many of us.

II. The problem of factorial analysis

The scores of each test being set off along a line, these can be given direction in space at angles whose cosines are their intercorrelations. They will then occupy a space of n dimensions if there are n tests, and the population of persons will be represented by points in this space, congregated round the origin where the man who is average in every test is situated. If standardized scores are used, the population will fall off in density equally in all directions from that point, its density contours being spheres.

The problem of factorial analysis is then to choose a set of axes or factors, preferably orthogonal or nearly so, to replace the tests as definers of the space, that is of the qualities of the persons. As there are innumerable sets available, some principles must be adopted to select one as the best. 'Best' may mean here best from some psychological, or from some mathematical, point of view, or possibly these may click into agreement in some system acceptable to psychologists and mathematicians alike. Among the principles which have been adopted are the following:[1]

(i) Representing the experimental facts (within the limits of the errors present) by a smaller number of factors than tests. There is probably unanimity on this point, but not on what the smaller number of factors are to be required to do best.

(ii) Reproducing as much as possible of the *whole variance* with each successive factor (Hotelling (6)).

(iii) Reproducing the *correlations* with the minimum number of common factors (Spearman (9), Thurstone).

(iv) Insisting on a *general factor g* (Spearman, Holzinger).

[1] I have placed after each, as a guide to the reader, the names of one or two of those most closely identified with the principle, but the classification is not of course perfect or complete. Kelley's name (7) will be missed. He is very important, but I have not felt able to identify him sufficiently with any of these principles. Stephenson's name can accompany Spearman's wherever it appears, and he is also interested in the relationship between test factors and person factors (10), though he does not make a principle of their reciprocity as Burt does.

(v) *Rotating* the factors until they become psychologically significant (Thurstone (11, 12)).

(vi) Requiring '*Simple Structure*', i.e. certain mathematical relations between factors and tests in a battery. These requirements have led to *no general factor* being found in tests reputed to be saturated with *g* (Thurstone (11, 12)).

(vii) Requiring *invariance of analysis* of the same test in different batteries, when used on equivalent samples of persons.

(viii) Using factors and loadings which are *reciprocal* for persons and tests (Burt (2, 3)).

III. Reproducing the variance

The natural desire of the statistician (e.g. Hotelling) is to lose as little information as possible. He adopts therefore the second principle, using the principal axes of the ellipsoids which result from the density spheres when the test lines are pulled into orthogonality (6). It will of course require as many factors as there are tests, to reproduce all the information. But if, after a few factors, the greater part of the variance has been accounted for, the remaining factors are neglected. The statistician has then arrived at the best shorthand description of *this* battery. He makes no claim to have arrived at factors which have any other significance, nor that they are invariant when tests are added to or taken from the battery. His axes indeed are not factors at all in the sense intended by those who use principles (v) and (vii).

There is, however, nothing to stop a worker who proceeded on these lines from then adopting principle (v), and rotating his few main factors in their own space in the hope of making them psychologically significant; but this has not in practice been done with 'principal components' based on the whole variance.

The first few factors found by the above method will approximate both to the total variance and also to the correlations. But it is the former they do well.

IV. Reproducing the correlations

(1) *Minimal rank methods*

Better approximation to the correlations can be made by another method, which depends on the fact that a matrix of correlation coefficients between mental tests, when suitable numbers are inserted in its diagonal, is of low rank, within the limits of error. The reason for this is, in my belief, the complexity and plasticity of the mind, and the operation of the

laws of chance. But whether this is the true reason or not is immaterial to our present argument, for the phenomenon is undoubtedly present and the result is that the correlations can be closely imitated by a small number of factors compared with the number of tests. This small number of factors does not however by any means give a good approximation to the whole variance but only to certain fractions called 'communalities'. There is therefore much information lost by using only these common factors. The other factors—specifics—are not needed to imitate the correlations of this particular battery. But they may be of importance in predicting success in some occupation. These minimal rank methods, therefore, which make much of the fact that they are parsimonious in common factors, attain that parsimony only by contenting themselves with imitating the correlations, and giving up any attempt to reproduce all the information contained in the tests. Their second great weakness is that since the whole number of factors exceeds the number of tests, each is subject to an indeterminacy and can be estimated with less accuracy than can one of the principal components of the former method.

(2) *The Spearman school*

Among those who use the minimal rank method there are two schools, that of Spearman (9), who discovered the fact of low reduced rank many years ago, and that of Thurstone (11). The method of the former school is to proceed by purifying a battery of intelligence tests until one general factor g is sufficient to account for the correlations. The g-saturations of the tests in this battery are then known. In more complex batteries this g is always taken out first. When a second factor is suspected, a battery is built up to measure it (by the residuals left after the removal of g), a battery which contains only these two. And so the process goes on step by step, a complex matrix of correlations being looked upon as made up by the superposition on top of the g-hierarchy, of sub-hierarchies each due to a new group factor. Invariance of analysis can be gained (principle (vii)) by retaining for any test that g-saturation which it had in the best 'g-battery' and so on with the other factors. No rotation is necessary.

Holzinger's Bifactor Method (4) is the same in principle, although it proceeds on more wholesale lines.

(3) *The Thurstone school*

In contrast with the step by step Spearman method, the Thurstone method is a simultaneous one. No attempt is made to reduce any battery by purification to rank one. Instead, a large battery is analysed simulta-

neously into a number of common factors—plus the specifics characteristic of both schools—by a method which involves first of all guessing the communalities and then taking at each stage the centroid of the residues. The resulting factors are devoid of psychological significance and have to be rotated in a search for positions where they will have such significance.

Here Thurstone introduces the idea of 'Simple Structure', a mathematically defined position of the axes (11). It is his belief that if in any battery 'Simple Structure' can be attained the factors will then be recognizable as psychological entities. 'Simple Structure' is marked by the absence of negative saturations and *the presence of many zero saturations*, for it is laid down that there must be at least one zero saturation in each test, and several zero saturations in each factor. This excludes a general factor, and Thurstone naturally therefore does not find one (12), even when he analyses tests said to be highly saturated with *g*. Just as naturally Holzinger and Harman (5), analysing the same data, do find a *g*, its presence in their pattern being, as they frankly say, "due to our hypothesis of its existence".

Alexander (1) has carried out an analysis using Thurstone's way of arriving at a set of minimal rank factors, but not his hypothesis of 'Simple Structure'. Instead, Alexander at this point passes over to the Spearman plan, and adopts for some of his tests analyses already known. This throws into relief the fact that the essential difference between Thurstone and Spearman is in the former's use of 'Simple Structure'. Spearman has no need, in his step by step method, to rotate his factors. But if he did use Thurstone's first analysis (as Alexander has done) he could then rotate to positions corresponding to the results of the step by step procedure. If, however, a battery of tests was put together from Spearman analyses, and offered for alternative analysis via 'Simple Structure', it would be necessary, to give a chance of agreement, to include several tests with nearly zero saturation in *g*, otherwise *either* 'Simple Structure' would not be attainable, *or* the two methods would be foreordained to disagree.

V. Reciprocity of tests and persons

The last principle (reciprocity of persons and tests) has been put forward in its rigorous form by Burt only (2, 3), although Stephenson (10) has proposed to check analyses of tests and of persons by one another. Burt holds that those factors are the proper ones to use of which it can be said that the same results are arrived at whether we analyse tests or

persons—the factors and loadings of the one analysis being the loadings and factors of the other. This end however Burt attains only by removing from the operation of the principle a factor consisting of the average performance of each person in all tests, and by working with unstandardized scores and covariances instead of correlations. Actually he removes his 'average' factor not by analysis but by selection of cases equal in average score, and it is not at all clear that he would obtain the reciprocity he desires were he to do this first for persons selected as equal and secondly for tests selected as equal.[1] On a matrix of scores which has somehow been centred both ways his reciprocity principle however does work. Such a matrix has only $n-1$ dimensions (one factor, the average, having been removed) and Burt uses the principal axes of its ellipsoids. His factors are therefore free from 'indeterminacy'. He does not however rotate them but interprets them as they are with their numerous negative loadings, a procedure which has passed muster in the analysis of temperament to which alone so far he has applied it, but would arouse opposition were it applied to the analysis of more ordinary mental test scores. His use of raw scores just as they stand is certainly wrong: but if a suitable set of units could be discovered—I have suggested one based on the sampling theory—the practice of analysing covariances instead of correlations would have much to commend it.

VI. Various other matters

There remain several matters which I have space only to mention. For a long time workers were content to analyse batteries of tests into factors, and did not proceed to the practical estimation of these factors in individuals. Now that this is beginning to be done, a realization is growing of how uncertain such estimates are, as a result of the mathematical weakness of postulating in minimal rank systems more factors than tests. It is also coming to be realized that to estimate factors, and from these to estimate proficiency in an occupation, is a roundabout procedure when the latter can be estimated direct.

Finally, much has been done to illustrate the influence on factors of selection among the persons, and also of training and maturing. In connexion with selection it has been made clear that factors are bound

[1] This was written in November 1938 and is, as I now think (February, 1939) too sceptical on my part. I would like to withdraw lines 33 to 35 on page 290 of my book *The Factorial Analysis of Human Ability*, for I now see that since Burt chose his subsample of persons to be not only equal in average to one another, but equal to the average of all, those lines are irrelevant.

to change, emerge, and disappear with the changing sample, and therefore also with the generations of man, and Price (8) has indicated how homogamy can influence the factorial make-up of successive generations. Factors are statistical coefficients, changing with the sample and the conditions and dependent upon stated assumptions: but with defined conditions and assumptions they are most useful as descriptive terms.

REFERENCES

(1) ALEXANDER, W. P. (1935). "Intelligence, concrete and abstract." *Brit. J. Psychol. Monogr. Suppl.* No. 19.

(2) BURT, C. (1937). "Correlation between persons." *Brit. J. Psychol.* XXVIII, 59–96.

(3) —— (1938). "The analysis of temperament." *Brit. J. med. Psychol.* XVII, 158–88.

(4) HOLZINGER, K. J. (1935). *Preliminary Reports of Spearman-Holzinger Unitary Trait Study.* No. 5. *Introduction to Bifactor Theory.* Chicago.

(5) HOLZINGER, K. J. & HARMAN, H. H. (1938). "Comparison of two factorial analyses." *Psychometrika,* III, 45–60.

(6) HOTELLING, H. (1933). "Analysis of a complex of statistical variables into principal components." *J. educ. Psychol.* XXIV, 417–41 and 498–520.

(7) KELLEY, T. L. (1935). *Essential Traits of Mental Life.* Harvard Univ. Press.

(8) PRICE, B. (1936). "Homogamy and the intercorrelations of capacity traits." *Ann. Eugen., Lond.,* VII, 22–7.

(9) SPEARMAN, C. (1927). *The Abilities of Man.* London: Macmillan.

(10) STEPHENSON, W. (1936). "The foundations of psychometry: four factor systems." *Psychometrika,* I, 195–209.

(11) THURSTONE, L. L. (1935). *The Vectors of Mind.* Univ. Chicago Press.

(12) —— (1938). "Primary mental abilities." *Psychometric Monogr.,* no. 1. Univ. Chicago Press.

THE FACTORIAL ANALYSIS OF ABILITY

II. DETERMINATION OF FACTORS[1]

By C. SPEARMAN

I. *Disappearance of 'G'* (p. 78).
II. *Grounds for determining factors* (pp. 78–80).
III. *Mechanism of producing zeros* (pp. 80–82).
IV. *Multitude of factors* (pp. 82–83).

I. Disappearance of 'G'

After the admirably broad and impartially expositive paper of Prof. Thomson, apologies seem needed for now descending to a matter of comparatively narrow scope and controversial status. But to some readers an intervention on my part will not be surprising, when they learn that the point at issue is nothing less than the existence of a general factor in ability; 'G', as it has been called.

Many a time already, no doubt, this factor has had to endure rude onslaughts. But hitherto all such attacks on it appear to have eventually weakened into mere attempts to explain it more simply. Now, however, there has arisen a very different crisis; in a recent study, nothing has been found to explain; the general factor has just vanished. Moreover, the said study is no ordinary one. Alike for eminence of the author, for judiciousness of plan, and for comprehensiveness of scope, it would be hard to find any match for the very recent work on *Primary Mental Abilities* by L. L. Thurstone. With great interest, accordingly, we proceed to inquire by what method he arrived at such a result, so opposed to all those gained previously, even by himself. To some extent this question has already been touched by Prof. Thomson. We will here take the liberty of treating it a little further.

II. Grounds for determining factors

Let us commence by considering the general operation of factorizing. Here, the initial data consist in the correlations between some selected variables, such as mental abilities; and the task is to infer the factors by which these correlations may be explained.

[1] A contribution to the Symposium presented at the Extended General Meeting of the British Psychological Society held at Reading in April 1939.

Now, the great difficulty in the inference is that the data are insufficient; as is well known, one and the same set of correlations can always be factorized in an infinity of different ways. On what grounds shall the choice be made between these?

Such preference appears to be regulated by two general principles; that of statistical simplicity, and that of psychological significance. Each principle may be applied in various manners.

As an outstanding instance, let us take the determination involved in the theory of Two Factors applied to abilities whose inter-correlations present the arrangement which has been called "hierarchical". Of the infinitely diverse possible factorizations of such a set of abilities, the said theory has selected that which derives all the correlations from one single factor. Statistically, this determination is grounded on its extreme simpleness. Psychologically, it is credited with affording the sole base for such useful concepts as those of 'general ability', or 'I.Q.'

But as soon as the abilities at issue begin to deviate from the hierarchical arrangement, then the reduction to a single factor becomes impossible, and so the whole situation changes. One might think, indeed, of replacing the reduction to a single factor by a reduction to as few as possible. But if this multiple reduction be adopted throughout—that is to say, without any additional restriction—then the determinateness is really forfeited. For whereas the reduction to a single factor can only be effected in one way, a reduction to two or more can be made in an unlimited number of ways.

The lost determinateness can, however, be more or less regained by stipulating that the weights of the respective factors should be heaped up in order of calculation; that is to say, the first factor is given the largest possible weight, the second factor as much as then still remains possible, and so on. Some such determination has been employed by Hotelling, Kelley and Burt. The trouble this time is not so much statistical but psychological; the procedure generates a large and regular number of negative values which contradict all experience.

The contrary may be said of a further ground that has been employed in the quest for determinateness. It was introduced, apparently, by Alexander and is now adopted by Thurstone himself. It consists in rejecting all factorizations which, like those just considered, introduce negative weights. Psychologically, as we have just seen, this rejection is well founded. But statistically the indeterminateness is not eliminated; only decreased.

Very dissimilar, however, is another ground which has been adopted

by Thurstone *additionally* to the preceding one. It is a new form of the statistical principle of simplicity. He writes: "the simplest order has been found when each row of the factorial matrix has at least one zero and when the number of zeros has been maximized."

And here at long last, not in anything we have seen hitherto, we encounter the cause of the unexpected disappearance of the general factor. Thurstone works in two stages. He starts off with his centroid procedure, so applied as to heap up the weights in order of calculation; so far the general factor is dominant, if not even exaggerated. But then he introduces his rotation of axes so as to: (*a*) eliminate all negative values, and (*b*) maximize the number of zeros. Here (*a*) is still harmless enough. But (*b*) is revolutionary; in particular, it alone is responsible for the general factor being eliminated.

III. Mechanism of producing zeros

Now, such a great and potent innovation in the method of determining factors would appear to deserve the keenest inquiry into its precise mechanism. Thurstone himself, however, tells us surprisingly few details about it. Departing from his usually irreproachable explicitness, he only remarks: "many different methods have been used in these successive rotations, and it would be impossible to describe them all."

In such default of further information, we must do what we can for ourselves. And for this purpose we will employ a few very simple examples.

Let the following figure represent a factorial matrix where the a's are the entries for four factors:

	α	β	γ	δ
Ability I	$a_{1\alpha}$	$a_{1\beta}$	$a_{1\gamma}$	$a_{1\delta}$
	.	.	.	.

As our first case, we will suppose the column headed by $a_{1\alpha}$ to have the largest possible total value; $a_{1\beta}$ the next largest possible; whilst $a_{1\gamma}$ and $a_{1\delta}$ approximate to zero.

Portraying the same absolute values in a profile we get Fig. 1. Here the weights of the four entries are given respectively by their ordinates α, β, γ, and δ.

This case, obviously, belongs to the type that we have characterized as that whereby the weights are heaped up as much as possible. To this extent it agrees with the result obtained by Thurstone in his first step.

But now these weights, whilst still retaining conformity with the

given correlations, admit of infinitely varying redistributions; this can be done, either by rotations of the axes, or otherwise. One such redistribution is that which may be termed a levelling-out. The ensuing profile is as in Fig. 2. And as far as the levelling is concerned, we now find an agreement with Thurstone's second step. For by that, too, the weights from first to last maintain a surprisingly uniform value.

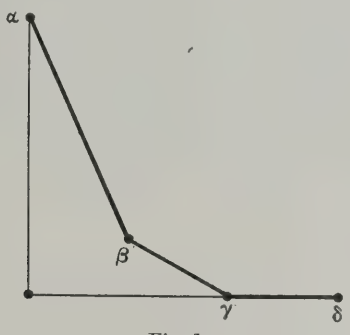


Fig. 1.

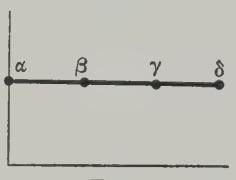


Fig. 2.

Here, however, the agreement of Fig. 2 with Thurstone's result breaks down. In the figure, there is no absence but rather an abundance of general factors. And the number of zeros, far from being maximized, diminishes from two to none.

Accordingly, to render this levelled-out profile more fully congruent with Thurstone's result, some further influence must be postulated. And this appears to be supplied in the fact that his method did not—and could not—demand the maximum number of exact zeros, but only of weights 'near' to this value. These are the weights which, since they fall below a certain minimum, are taken to be zero in the sense of 'not significant'; they are ascribed to mere errors (in large measure, those of sampling).

But all this suggests a startling consideration. What happens, we must ask, when the weights are very small throughout whereas the errors are very large? The eventual result must be that all the weights become less than their errors, so that they are all counted as 'zeros.' In general, then, we see that very numerous zero loadings may after all be merely the effect of very large experimental errors. And just this unfortunate issue would seem to have befallen Thurstone.

Accordingly, we cannot well escape from the suggestion that his otherwise eminently admirable work has what the French call "the defects of its qualities". In the first place, the extraordinary comprehensiveness of his tests left but scant time for each of them.

Hereby the correlations have necessarily been diminished. And secondly this same multiplicity of tests rendered the labour of calculating the ordinary coefficients of correlation prohibitive. He had to fall back on the tetrachoric coefficients, whereby the errors are trebled.

IV. Multitude of factors

So far, out of all the treatments according to which the factor-weights could be distributed at will and so rendered determinate, we have only considered two special types; they are those of heaping-up and of levelling-out. Yet a third type also must now be brought into consideration. It may be called that of scattering. For here the weights, without altering their pattern otherwise, are scattered among more numerous factors.

If now this additional treatment be applied to our profiles, their likeness to the results of Thurstone becomes still further increased. This is true both of his initial heaping-up and of his final levelling-out. These two results of his, although so widely different otherwise, nevertheless agree in that the factors introduced are extremely numerous; no less than twelve.

Now, there are grounds for regarding this multitude with suspicion. I have myself shown elsewhere that all his results can be just as well explained with only half the number[1]. Moreover, he concedes that for five out of the twelve—although quite as large as the others—even he himself cannot suggest any psychological meaning. Yet again, we may note that when one and the same total weight is divided up into more numerous portions, these must individually become smaller, and therefore more than ever submerged in error. And so we begin to fear that the numerousness of the factors is after all only due to the arbitrary postulate of numerousness of zeros.

Anyway, the introduction of multitudinous factors certainly has an all-important consequence. However large a general factor may really be present, the share of it devolving on any single one of the divided-up factors must tend to become unnoticeable. Statistically, it will be 'insignificant'. And so at last the whole mystery of the vanishing general factor would appear to be explained. All the available evidence consistently supports the view that here this factor only ceases to be visible because of being divided up into very small pieces, which are then submerged in large errors.

[1] "Thurstone's work re-worked." *J. Ed. Psychol.* (1939), xxx, 1–16.

By what further information such cases of its mere submersion can be still more crucially distinguished from those of its genuine absence—this is a nice problem which must be reserved for a later occasion.

Anyway, the whole course of the preceding considerations has emphasized the fact that *no* hard and fast procedure of factorizing is adequate. Since the correlations fail to determine the factors completely, some further grounds are needed for this purpose. And only in one special case, apparently, is any such ground ready to hand. It is the case of reducibility to a single factor. Here the ground is the very cogent one of parsimony. In all other cases, to find adequate grounds is a matter of no small difficulty; every case must be treated on its own merits.

THE FACTORIAL ANALYSIS OF ABILITY

III. LINES OF POSSIBLE RECONCILEMENT[1]

By CYRIL BURT

I. *The meaning of factors* (pp. 84–85).
II. *The sampling theory and the two-factor theory* (pp. 85–86).
III. *Interpretations of G* (pp. 86–89).
(*a*) *Sampling the neural bonds* (pp. 87–88).
(*b*) *Sampling the tests* (pp. 88–89).
IV. *The reciprocity principle* (pp. 90–92).
(*a*) *Variations in variance* (p. 90).
(*b*) *Negative saturations* (pp. 90–92).
V. *Special issues* (pp. 92–93).
References (p. 93).

Prof. Thomson, in a paper as concise as it is impartial, has covered almost the whole field of factor analysis. Prof. Spearman in his defence of the general factor and Dr Stephenson in his paper on *Q*-technique have raised problems more specialized and distinct. It has been suggested that I should deal mainly with two fundamental issues: (i) the sampling theory as opposed to the two-factor theory, and (ii) the reciprocity of tests and persons. Even so, in the space allowed, I can touch only upon one or two outstanding questions, with a curtness they scarcely deserve.

I. The meaning of factors

Much apparent disagreement would vanish if we all kept to the primary meaning of the term factor. The matrices that define factors in psychology have the essential properties of 'selective operators'. Hence a factor is primarily a principle of classification and nothing more: it is expressed in quantitative form simply because the items, whose characteristic pattern constitutes the distinguishing mark of the class described by it, vary continuously and in degree rather than discontinuously and in kind. Now, just as the traditional logic of classes has recognized four kinds of possible predicable—genus, species or differentiae, individual propria, and accidents—so factor analysis is led to

[1] A contribution to the symposium presented at the Extended General Meeting of the British Psychological Society held at Reading in April 1939.

envisage four kinds of possible factors—general, group or difference, specific, and accidental ((6), p. 259). The two-factor, three-factor, and multiple-factor theories are all special derivatives of this four-factor theory. And most of the difficulties and controversies that trouble the factorist in psychology have already been debated at great length by different schools of logic. The temptation to identify an abstract and hypothetical factor, reached by mere mathematical analysis, with a real and concrete 'factor in the mind' (e.g. with an 'ability') is simply a new instance of the old problem that has divided nominalists, realists, and conceptualists for centuries. So long as we regard the factors reached by Spearman and Thomson as descriptive components only, and not as causal entities, there is (as I hope to show) no serious conflict between their theories.

II. THE SAMPLING THEORY AND THE TWO-FACTOR THEORY

Thomson's sampling theory, like Spearman's two-factor theory, was originally put forward to explain the 'hierarchical order', which Spearman was the first to notice, among the coefficients obtained by correlating cognitive tests. The sampling theory maintains that "the hierarchical order is the natural relationship among these coefficients". I agree; but on more general grounds. By a hierarchy I assume we mean a matrix of rank one.[1] Now, as I have elsewhere shown ((7), p. 187), the process of self-multiplication, if sufficiently repeated, will ultimately reduce *any* matrix to a matrix of rank one. But the process of correlating is essentially a process of self-multiplication. Hence to correlate is to take the first term in a converging sequence of matrix products which must eventually end in a hierarchy as perfect as can be desired.

To my mind, therefore, the question is not: Why are our correlation tables hierarchical, i.e. why have they only one latent root, so that the reduction seems virtually completed in a single step by the sheer process of correlating? Such cases were nearly always the illusory effects of high probable errors. The question should rather be put: Why is the first latent root so much larger than the second, and *a fortiori* larger than the rest? My answer is: Because of the way in which the psychologist has selected his tests. Examine the correlation tables cited as evidence of

[1] Not all the hypothetical tables given by Spearman and his pupils to illustrate a perfect hierarchy are matrices of rank one: cf. (10), vol. II, p. 274, and (3), p. 175 (where "it is assumed that in a hierarchy the coefficients decrease *arithmetically*", and *not* proportionately: e.g. in the table given as an example of a perfect hierarchy "the correlations in both rows and columns decrease throughout by 0·05").

a marked hierarchical tendency: it will usually be found (i) that the tests were all chosen from one generic type, in most cases cognitive; (ii) that only positive examples of the type were taken, i.e. the tests always measured ability, never disability; further, that the tests were picked to represent, as far as possible, (iii) widely different subtypes, and (iv) widely different levels of complexity; but that, at the same time, (v) each subtype, and usually each level, was represented in most researches by one test only. These conditions of themselves are sufficient to ensure several striking features in the resulting correlation tables: we may safely infer that the correlations will be almost always positive and never (except for the accidents of sampling) negative; will range from small to high values; tend to obey the proportionality criterion, but also tend to exhibit a slight, diminishing and (as a rule) cyclic overlap. Now the conditions specified cover the essential requirements of 'random' selection. Thus the hierarchical tendency arises, not so much because the tests have blindly sampled the bonds at random, but *because the psychologist has systematically sampled the tests at random.*[1]

III. Interpretations of G

There is no mathematical incompatibility between the two-factor theory and the sampling theory, as their respective advocates at one time maintained. Applied to the facts that both are concerned to interpret—the simple hierarchical tendency—the two procedures give virtually the same result. As I have tried to demonstrate by algebraic proof, if we estimate for every person the total of the test measurements to be expected on Thomson's sampling theory, those totals will possess all the properties of a general factor. Conversely, as Thomson has shown, by an ingenious transformation matrix[2] a two-factor matrix can always be rotated to give a sampling theory matrix.

The real question, therefore, is—which theory best fits what we know of the working of the mind or central nervous system? And this (in spite of Stephenson's protest against "irrelevant attempts to bring neural interpretations into the picture") is an issue that can be answered only by non-mathematical evidence. As Spearman himself has observed: "the burden of further proof is in large measure transferred to physiology".

[1] Lack of space compels me to be content with unqualified and undemonstrated statements. But the interpretations put forward in §§ I and II will be at once clearer and more plausible if the reader thinks of what would occur with correlations, not between tests or traits, but between persons.

[2] The transformation matrix as it stands is non-singular. Consequently Thomson may not accept the previous statement.

(*a*) *Sampling the neural bonds*

Stephenson has maintained that "there are in the human being... only a few unitary traits, born out of system (1) analysis", i.e. "out of the analysis of primary abilities by *r*-technique with two-factor methods". Thomson, on the other hand, maintains that "far from being divided up into a few 'unitary factors', the mind is a rich, comparatively undifferentiated complex of innumerable influences—on the physiological side an intricate network of possibilities of intercommunication". The latter description is evidently in much closer accord with the neurologist's conception.

Spearman would identify a man's *g* with his mental 'energy' (should it not be 'power'?). Thomson favours the view that "the general energy, if there be such a thing, is...of an instinctive and emotional nature rather than intellective". With this, of course, I agree, since I identify 'mental energy' with 'general emotionality', not with general intelligence; the latter (as McDougall and Maxwell Garnett have argued) *directs* the general energy. Thomson's own theory, however, seems to exclude all possibility of direction.

He begins by arguing that the correlations between cognitive tests can best be explained, not by attributing them to the amount of mental energy each man has, but rather to the number of neural connexions each man can organize. This I readily accept, so far as it goes. Moreover, his further postulate of 'pools' and 'sub-pools' seems to me to be quite in keeping with that combination of general homogeneity of cellular structure with minor differentiations in different parts (i.e. in different layers and areas of the cortex), and of mass action with partial localization, that the structure and functions of the brain unquestionably exhibit. But when we ask how the different tests succeed in setting this 'structural network' into operation, Thomson's account appears somewhat inadequate. I willingly admit that a blind, trial-and-error process is set up by many of our tests (especially by tests blindly and therefore often inappropriately selected by the tester—e.g. some of the tests that have been applied to duller creatures, whether animals or human beings); but I should maintain that the higher tests of intelligence as applied to higher minds operate specifically by *selecting* the 'bonds' rather than by *sampling* them. Here, therefore, I agree with Spearman that a 'good test of *g*' will not stimulate neural connexions at random, but is designed to educe the particular connexions that are relevant to the test problem.

I should even go further: the eliciting of blind associations and the eduction of relevant relations are to my mind only partial aspects of the essential function of the mind or nervous system, namely, its *integrative action*. From this standpoint, as it seems to me, both theories can be reconciled.

These two theories, however, envisage only correlations between tests, not correlations between persons, and only tests of cognitive abilities, not tests or assessments for temperamental traits. But, as recent work has amply shown (e.g. (11)), when we extend the scope of our experiments to include these newer problems and methods, the hierarchical tendency still predominates in the resulting tables: or, to speak more accurately, the latent roots of the correlation matrices continue to show much the same rapidly diminishing proportions in both cases (approximately 50, 16 and 5% of the total variance, when no special selection operates). This again suggests that the hierarchical tendency (or what is loosely described as such) may be the direct effect of the process of correlating; but it does not prevent us from postulating some feature in the mind which allows this effect to manifest itself undistorted. Would Thomson accept the energy hypothesis here, to explain the positive and proportionate correlations between emotional traits? Or would he account for the so-called 'energy factor', and the other general factors that are sometimes found, by the partial overlapping of numerous elementary tendencies? When we consider the apparent nature of the factors so far revealed, it would seem more plausible to attribute the secondary factors at any rate, not to the random sampling of the elementary tendencies, but rather to the way those tendencies are organized, i.e. to certain *general attitudes or dispositions, characteristic of the whole personality*. If so, the integrative hypothesis will cover both fields.

(*b*) *Sampling the tests*

Elsewhere (8) I have pointed out that, just as in correlating tests we have to consider how we have selected or sampled our persons, so in correlating persons we are compelled to consider the fact that we are now either selecting or sampling the tests or traits; and much the same principle applies in the second case as in the first: i.e. the correlations between persons can be most readily interpreted if the "tests form a random sample of the total population of tests" ((8), p. 68). If we deduce the natural consequences of this condition, we are led to theorems which closely resemble Thomson's sampling theorems and his related theorems

on selection, except that our propositions now refer, not to the fact that our tests are sampling or selecting the mind's neural bonds, but to the fact that *the psychologist is sampling or selecting the tests or traits*. For example, just as the factor measurements for the persons obtained by correlating *tests* measure the complexity (or 'richness', as Thomson calls it) of those persons' minds, so we might say that the factor measurements for the tests obtained by correlating *persons* measure the complexity or 'richness' of the tests. Now, Thomson has pointed out that, when correlating tests, the richness of the tests is measured by their saturation coefficients. But, if I am right in the general equivalences that I have endeavoured to establish between the two methods of approach, then, by the so-called 'reciprocity principle', saturation coefficients for the tests, obtained by correlating persons, are virtually identical with the factor measurements for the tests, obtained by correlating tests. This seems to bring my own speculations closely into line with Thomson's views.

No doubt, these deductions from the wider theory of sampling will not throw much light on the analysis of the mind; and it is rather with this latter problem that Thomson, like most psychologists, is more immediately concerned. But they certainly throw light on the important practical problems that arise in actual examining and testing (e.g. when we consider that any particular examination paper can include only a limited sample of the entire 'population' of questions that could be asked about a given subject).

On one minor point I differ both from Thomson and from Spearman. From Thomson's sampling theory it follows that, on a *single* application, a test of very slight complexity or 'richness' would give low intercorrelations and low self-correlations (reliability coefficients), but that, if repeated and corrected for unreliability, it would give as high a correlation with other tests (and therefore with the general factor) as the most complex. Spearman has similarly maintained that, after due correction, there exists "a correspondence between what may provisionally be called General Discrimination and General Intelligence which works out with great approximation to one or absoluteness" ((1), p. 284). Now in my own early experiments I endeavoured to include (if I may borrow Spearman's terminology) tests of "all forms of sensory discrimination" and also of "the more complicated intellectual activities". I found that simple sensory and simple motor tests often gave high reliabilities, and yet, even when corrected, furnished comparatively small correlations with cognitive tests of a more complex type. This seems to me to fit much more closely with a 'complexity' (or integrative) theory of intelligence than with either a sampling theory or an energy theory.

IV. The reciprocity principle

May I now turn to the criticisms advanced by Thomson and Stephenson against my own conclusion, namely, that the non-general factors obtained by correlating persons are the same (except for minor differences of weighting) as those obtained by correlating tests?

(*a*) *Variations in variance*

Stephenson's chief objection to my reciprocity principle, as he terms it, is that it implies objective differences in variance. Variance, he insists, "depends solely on the whim of the psychologist"; hence, the first step in any form of factor analysis should be to convert the raw measurements into standardized scores, and treat the variances as equal. Now anyone who has investigated individual variation for different capacities and attainments (e.g. to compare it with the variation between different ages or different sexes) quickly learns that it is by no means dependent solely upon the psychologist's whim. As I have elsewhere pointed out, there is a vast accumulation of evidence indicating that *variance depends largely, if not mainly, on the complexity of the process measured.* Indeed, the method of factor analysis that I have found most successful entails the identification of the estimated self-correlations, not with unity (as Hotelling has assumed) nor with the highest correlations for each test (as Thurstone proposes), but with a proportionate estimate of this differing variance. The implication which Stephenson criticizes has therefore always formed in my view a merit of the reciprocity principle, not a defect.

(*b*) *Negative saturations*

But the strongest criticisms have been directed against my acceptance of 'negative loadings just as they are'. As Spearman puts it, the 'negative values' for the saturation coefficients 'contradict experience': and with this both Thomson and Stephenson, as well as Thurstone and Alexander, would seem to agree. Factors, it is assumed, are 'abilities in the mind'; and to talk of an 'ability' as something which *hinders* a person from performing a test successfully seems a contradiction in terms. My defence is twofold: the assumption tacitly attributes to the 'factors' a meaning which is at once (i) too intellectualistic and (ii) too concrete.

(i) Performance in any test is a result of conative energies as well as of cognitive capacities. A child who devotes most of his time and interest to one subject must give less time and interest to others. Conversely, lack

of ability in one subject may, by compensation, lead to added effort in another. Moreover, as experiments on transfer amply show, the acquisition of one kind of ability often interferes with the acquisition of a second. And generally, we may ask, when reciprocal inhibition and occlusion are such well established features of neural activity, as mental conflict is of higher mental processes, why should factor analysis, for purely *a priori* reasons, refuse all recognition to interference-factors?

(ii) But the negative saturations may be defended on a more fundamental ground. Strictly regarded, test measurements do not describe minds; they describe relations between minds and their environments; and the factors that result denote, as we have seen, abstract principles for classifying such relations. Now classification may proceed by logical dichotomy just as well as by positive and co-ordinate classes. Bipolar factors classify by antitheses—subjects, for example, into verbal subjects and non-verbal, non-verbal subjects into arithmetical subjects and non-arithmetical, and so on: group factors classify by co-ordinate classes—verbal, arithmetical, manual, and the like. Or, if we prefer to apply the classification to persons, the same factors will lead us to classify pupils in the same alternative ways. It is the negative saturations that arise in this second application—i.e. the negative saturations deduced by correlating persons—that appear to cause greatest difficulty to my critics: for example, in experiments on imagery, if a person is a good visualizer, the saturation coefficient for my first bipolar factor gives him a positive mark; if he is bad, a negative. Yet surely this is nothing but the common device of translating absolute measures into deviations about an average. If abilities can never be expressed by negative figures, how does Stephenson defend the negative deviations that inevitably arise when he converts his test measurements into standard measure?

The foregoing criticisms turn mainly on different views as to method. How do they affect the results? When I first proposed using the device of correlating persons for researches in which we were jointly interested, Stephenson was not alone in criticizing the assumptions which the method implied. But he went further. He put forward, among other fruitful suggestions, an alternative two-factor technique, which, he considered, would escape many of the objections urged against my 'multifactor technique'—a '*Q*-technique' to replace the so-called '*P*-technique'. With this modified method, he now believes, the factors that result from correlating persons, so far from being merely 'the obverse' of those that result from correlating traits (as he at first was inclined to agree), prove to be 'poles apart'. Working side by side in the same laboratory we have

since been able to apply our respective methods to much the same problems and often to the same material. The reader can, therefore, examine for himself the differences in our published results. There will, I think, be little doubt about the conclusion. In every field—scholastic preferences, aesthetic appreciation, temperamental types—the factors described by Stephenson are virtually[1] the same as those obtained by the earlier and alternative techniques.

V. Special issues

In reply to my criticisms of his first paper (on "*Q*-technique") Stephenson has further developed his views in a second paper (on "Abilities defined as Non-Fractional Factors"), and suggests that the distinctions between 'mathematical' and 'psychological' factors and between concrete 'abilities' ('non-fractional' factors) and abstract 'components' ('fractional' factors) will go far to meet my points. To discuss the way he redefines these distinctions and classifies cognitive processes in his table would require more than my allotted space. I will merely say, first, that I am delighted to find him now explicitly accepting the 'theorem of four factors'—i.e. four *kinds* of factors (general, common or group, specific, and error or chance factors);[2] and, secondly, that I readily accept his algebraic demonstration, since it would appear to be virtually the same as my own ((9), p. 167). On the other hand, I doubt some of the deductions that he draws, and am surprised by certain omissions from the arrangement of levels and subtypes in his scheme of

[1] The only large difference between our figures appears in the field which happened to form the subject of Stephenson's first experiment on *Correlating Persons instead of Tests*, namely, 'artistic taste' ((5), p. 21). Here, as one of my critics points out, Stephenson's highest correlation was only 0·66, whereas in a previous account I had given an average correlation of 'nearly 0·90' ((4), p. 280). But my figure was obtained with pictures deliberately chosen so as to differ as widely as possible in artistic merit, whereas Stephenson's was obtained with very homogeneous material—namely, simple colours: indeed, his findings should be compared, not with my discussion of the general factor of artistic taste (as shown in the experiments with pictures), but rather with my description of *types* of artistic appreciation (as shown by my own experiments with simple colours (4), pp. 280 *et seq.*). I may, however, add that in correlating persons (as was pointed out in my early work with Binet tests) the coefficients obtained are apt to look surprisingly high; and in Stephenson's more recent work the figures are much higher and approach far more closely to my own (cf. the summary in (11), p. 413).

[2] In the revised version of this passage the 'error factors' are apparently included in the 'specific', and not regarded as a separate kind of factor. It should be added that both Stephenson's discussions and my own have here been greatly abridged owing to the exigencies of space. But we have drawn up jointly a more detailed statement on the points over which we agree and disagree; and this will shortly appear in *Psychometrika*.

test construction.[1] But it is plain that we are now in much closer agreement than appeared at the outset of our discussion.

With Prof. Spearman's criticisms of Thurstone, I am largely in agreement, and can give independent confirmation of his chief conclusion. As soon as Thurstone's tables were available, Mr H. J. Eysenck subjected them to a factor analysis by the 'group-factor method' used in my earlier work (12). He finds (i) practically the same values for the general factor as Prof. Spearman, and (ii) practically the same non-general factors as Thurstone. Thus, once again it seems possible to reconcile views that at first appeared entirely opposed.

REFERENCES

(1) SPEARMAN, C. (1904). *Amer. J. Psychol.* XV, 201–93.
(2) BURT, C. (1909). *Brit. J. Psychol.* III, 94–178.
(3) CAREY, N. (1916). *Brit. J. Psychol.* VIII, 175.
(4) BURT, C. (1933). *The Psychology of Art*, ap. Burt, C. *et al. How the Mind Works.* London: Allen and Unwin.
(5) STEPHENSON, W. (1935). *Char. & Pers.* IV, 17–24.
(6) HARTOG, P., RHODES, E. C. & BURT, C. (1936). *Marks of Examiners.* London: Macmillan and Co.
(7) BURT, C. (1937). *Brit. J. educ. Psychol.* VII, 172–95.
(8) —— (1937). *Brit. J. Psychol.* XXVIII, 59–94.
(9) —— (1938). *Psychometrika*, III, 151–67.
(10) SPEARMAN, C. (1938). *Psychology Down the Ages.* London: Macmillan and Co.
(11) DAVIES, M. (1939). *Brit. J. Psychol.* XXIX, 404–21.
(12) EYSENCK, H. J. (1939). *Brit. J. educ. Psychol.* (in the Press).

[1] My own 'scheme of test construction' includes, on the cognitive side, a motor group as well as sensory groups, and associative groups (memory, habit formation) as well as relational: cf. for example, the scheme set out in (2), pp. 98–9, and later papers. But, as the earlier paper shows, I should willingly agree with Stephenson that the tests selected for factor analysis should not be (as he puts it) "any old things we care to throw together". Such a principle (or lack of principle) certainly suggests that the investigators were merely taking random samples of all possible tests; and the features of their correlation tables might therefore be plausibly regarded as the effects of such sampling. To some extent no doubt that is true. In practice, however, sampling that is not consciously controlled is not so truly random as might be supposed.

THE FACTORIAL ANALYSIS OF ABILITY

IV. ABILITIES DEFINED AS NON-FRACTIONAL FACTORS[1]

By W. STEPHENSON

I. *Introduction* (pp. 94–95).
II. *The relations between test construction and factor theorems* (pp. 95–99).
(1) *Spearman factors* (pp. 95–96).
(2) *The universal factor,* G_u (p. 96).
(3) *The sense factors* (pp. 97–98).
(4) *The material factors* (pp. 98).
(5) *Relation and form factors* (pp. 98–99).
III. *Factor theorems* (pp. 99–100).
IV. *Abilities* (p. 100).
V. *Factors* (pp. 101–103).
(1) *Non-fractional factors* (p. 101).
(2) *Relation between nf and f factors* (pp. 102–103).
VI. *Conclusion* (p. 103).
References (pp. 103–104).

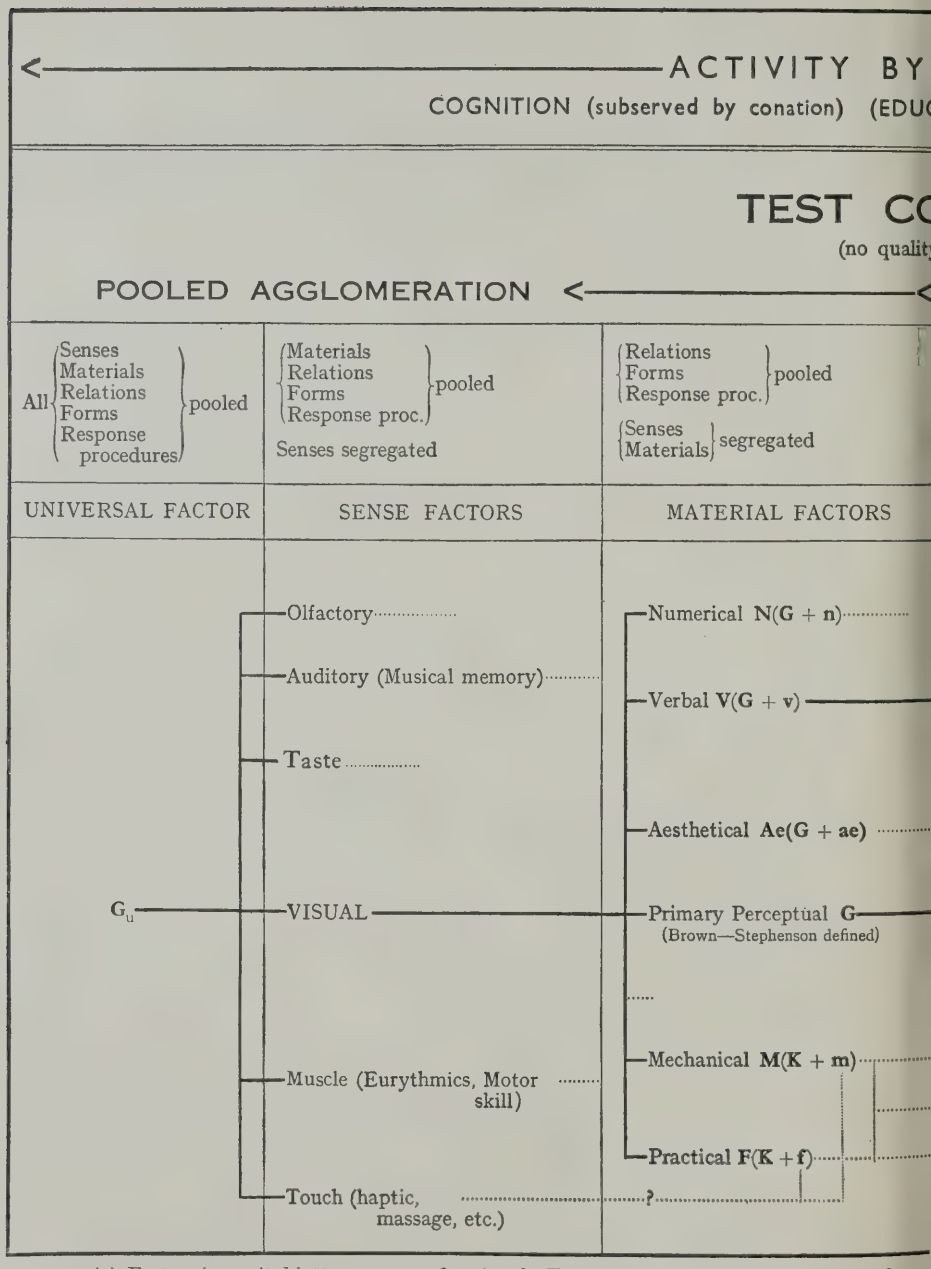
I. Introduction

The factorist, as psychologist, should set out to put questions to human nature. Apparently we have all very different questions to ask. Thomson asks questions about a vast nebulous sea, a protean mass, of innumerable brain elements; Burt tries to bring integration into this picture; whereas Spearman and I want to know whether memory is the same functionally as eduction, or whether abstraction is anyway different from generalization. Even as I say this it seems obvious who are asking psychological questions and who are not.

Again, except in Spearman work, I know of no factor systems in which some reference is made to the characteristics of the tests we use. Yet, so it seems to me, unless factor theorems refer in some way to the apparent characteristics of the variables we use, factorists cannot hope to speak the same language. As for the current factor methods, all the beauty and accuracy of Hotelling's procedures cannot hide their incompetency

[1] A contribution to the Symposium presented at the Extended General Meeting of the British Psychological Society held at Reading in April 1939.

Table I. *Showing connex*



(*a*) Factors in capital letters are non-fractional. Factors in small letters are fractional.

(*b*) The more uniform a test is, the greater the complexity of the *nf* factor representing it.

(*c*) Spearman's universal *g* factor has so far been defined in terms of an *nf* factor for primarily perceptual tests.

DIVIDUALS ——————————————>

tion, thinking, abstraction, attention, etc.)

UCTION

alance)

<—————————— UNIFORMITY

oc.} pooled gregated	{Response proc.} pooled {Senses, Materials, Relations, Forms} segregated	No pooling All {Senses, Materials, Relations, Forms, Response proc.} segregated
TION FACTORS	FORM FACTORS	TEST FACTORS
ical $P_s(V + p_s)$ .....) l $R(V + r)$ ..... g) V c.	Analogy $A(V + a)$ Classification $C(V + c)$ ..... Series $S_e(V + s_e)$ ..... Rhythm $H(V + h)$	Test S_1 $(A + t_1)$ Test S_2 $(A + t_2)$ Test S_3 $(A + t_3)$ ....
$(G + k)$	Analogy $A'(G + a)$ Classification $C'(G + c)$ ..... Series $S'_e(G + s_e)$ ..... Rhythm $H'(G + h)$	Test S'_1 $(A' + t_1)$ Test S'_2 $(A' + t_2)$ Test S'_3 $(A' + t_3)$ etc. ....

') Abilities (represented by *nf* factors) can be defined at any stage of test agglomeration.

) 'Capacities', 'powers', etc., should be independent critically of test construction, i.e. of material, relation, form, or test factors.

to deal with psychological problems. Thurstone's factors, likewise, are in no way defined in terms of psychological needs. Prof. Burt's reciprocity procedure, a most notable advance in technique, is insecurely founded, for as yet he has put forward no generally acceptable system of units, which is surely of first importance in any science.

I find very little help from these various methods for the concrete psychological problems that we seek to solve, and that, in the last resort, can only be solved by systematic factor studies. Our concern, in the major part of factor work, is with the activity of individuals, the tests we construct, and factor theorems. The way in which these impinge upon one another is worth illustrating, I think, in terms of Spearman factor studies.

II. THE RELATIONS BETWEEN TEST CONSTRUCTION AND FACTOR THEOREMS

(1) *Spearman factors*

I propose to consider what are the connexions between test construction and factor theorems, and how both relate to the activities that the tests set in motion in the minds and bodies of the persons we test; I do this by referring to actual work done on, or implicit in, Spearman's Theory of Two Factors.

Table I shows a broad landscape of the components in question. Only a few of the possibilities are given in the table, which reminds us that cognitive activity is under consideration. From work already done we exclude memory, speed preference, and sense acuities from the activity in question. This leaves perception, thinking, generalization, abstraction, conceptualization and the like, and eductive processes in general. Spearman's theory applies to *any* tests whatsoever for such activities, provided the tests are sufficiently dissimilar. Any such tests define Spearman's universal *g*-factor (and not, as Thomson supposes, merely the tests used by Brown and myself(1)). Spearman's *s* factors are at the outset merely that part of the variance not accounted for by *g* factor. But these *s* factors can be subdivided and made cogent, and an essential objective in Spearman work is that of determining just what subdivisions are permissible in fact and pertinent in psychology. Table I offers some information about such subdivisions.

Instead of regarding tests as any odd things we care to throw together, Table I seeks to bring a sense of orderliness into their construction. Tests are constructed as agglomerates or pools of test-units, and the pooling

takes place systematically in terms of test-units of different senses, modes of material, relations, forms, and response procedures. In Spearman work we distinguish between psychological factors such as the universal *g*, group factors, and specificity factors, and mathematical factors (general, common, and specific). But Table I goes further than this and tries to give some idea of the nature of certain psychological factors—instead of group factors[1] it shows *material* and *relation* factors, and instead of specificity (or minor group) factors it shows *form* and *test* factors.

(2) *The universal factor,* G_u

Systematic psychological work consists of constructing tests which endeavour as far as possible to deal with each mode of activity separately —that is, test systems can be produced in which abstraction is held critical, and these compared with others in which generalization is held critical. The activities are dealt with as variables to be controlled. But to begin with we could consider test-units for all forms of activity referred to at the head of the table. In this sense our concern will be with 'general' intelligence.

In the first column we contemplate tests for complete agglomeration —they are pools of all possible varieties of test-units, for all possible cognitive activities, senses, materials, relations, forms and response procedures. Such tests would presumably measure 'general' ability—*by which I would mean an ability in no critical way depending upon senses, materials, forms, relations, etc.* The agglomeration is no mere pool of *everything* pertaining to cognitive activity: it is devised so as to ensure no critical dependency upon the idiosyncracies of the senses, modes, forms, etc.

The variance for such tests would be divisible into

$$G_u^2 + s^2 = 1.$$

The *s* factor would be error, *and there could be no tangible sub-division of this into parts, except errors,* for we could not identify such parts with anything meaningful in psychology.[2] G_u is far the most comprehensive mixture of test-units that is conceivable. So far no such tests have been constructed, and no such factor isolated.

[1] These have always been regarded as not only those which extend into several variables, but those which, in addition, are of major psychological importance. It is easy to produce hundreds of *common* factors, but very difficult to arrive at even a handful of *group* factors.

[2] But when we subdivide the *activities* into different modes, these *s* factors can take on meaning with reference to such activities.

(3) *The sense factors*

The first subdivision would ordinarily be into the *sense* modalities involved. These would supply tests with test-units all of one *sense*, but with all possible activities, materials, relations, forms and response procedures pooled. These tests must not be confused with any for measuring sense acuities. On the contrary they exclude acuities. The tests contemplated are broader than any extant to-day, not less. Thus we might construct a hodge-podge test in which the fundaments all involve olfaction—our tests would be bottles of smells, etc., for different response procedures (selecting one bottle, perhaps, from a set of four), for different test *forms* (classification, series, etc.), for different psychological relations (especially likeness, psychological, and spatial).

Suppose a number of such agglomerate tests are made. The comprehensive pooling ensures as far as possible (theoretically) that materials, forms, relations and response procedures are not critical, and the following Spearman factors become possible:

nf factor*	Divisible into	*f* factors*
(*a*) Olfactory *O*	,,	$g^2+o^2+s_o^2=1$
(*b*) Auditory *A*	,,	$g^2+a^2+s_a^2=1$
(*c*) Visual *I*	,,	$g^2+i^2+s_i^2=1$
(*d*) Muscle *U*	,,	$g^2+u^2+s_u^2=1$
(*e*) Touch *W*	,,	$g^2+w^2+s_w^2=1$

* For explanation, see later.

Because of the gross pooling, the *s* factors could not be further subdivided meaningfully. The group factors *o*, *a*, *i*, *u*, and *w* would be particular to the respective senses. I do not say that all of these will exist (for no work has yet been done with such factors)—but any that did, could be identified in some way with sense modalities.

How would the *g* component be defined? From the internal evidence of the tests for sense segregations we might find it convenient or permissible to define it as a factor for one of the senses. For several reasons the visual sense suggests itself. Tests for visual modality would exclude symbolization and reading, in which memory, auditory and motor processes mediate. The tests would be composed of primary perceptual material—shapes, lines, colours, volumes, surface structures, etc. They would in fact be the same as those defining the material *G* factor (column

three of Table I). The variance for the visual sense *nf* factors would then be

$$G^2 + s_i^2 = 1;$$

and all the other sense-factors, *o*, *a*, *u* and *w*, etc., would be relative to this *G*.

(4) *The material factors*

Spearman work hitherto has centred largely on material factors. The major abilities discussed in popular psychology also come in this category. Thorndike's social, concrete, and abstract intelligences are concepts that merely embrace the materials each is concerned with—so are mechanical, artistic, practical, and scholastic abilities. All are inextricably bound up with particular kinds of material or psychological 'objects'. Table I lists factors *N*, *V*, *Ae*, *G*, *M* and *F*.

G has been identified with Spearman's universal *g* factor of the theory. It is the simplest thing to accept, and is in keeping with sound psychological notions. The other material factors are all divisible into *G* plus a group fractional factor. The factor *G* now plays two parts—it can be tested separately with visual perceptual tests, and also enters into all other tests. Such are the 'unitary' factors sought for by factorists. Consider, for example, *V* ability. Any verbal test, however narrowly segregated for form, relations, etc., will be divisible into *G* and *v*. But for sane definition of such a *v* component it would be wise to recognize pooling again: the tests of *V* ability would be agglomerates of all possible verbal relations, forms and response procedures. Can anyone doubt their efficacy when so defined? Their variance would be

$$G^2 + v_i^2 + s^2 = 1.$$

Again, because of the pooling, the factor *s* could not be meaningfully divided into smaller components. The factor v_i would tend not to be the same as that issuing from my earlier studies on *v* factor (6). It would be a more comprehensive fractional component.

(5) *Relation and form factors*

So we could continue, with increasing separation of the material into that of particular relations, forms, and response procedures. I am perhaps the only factorist who has tried to attend to relation factors, through the work of Wedek (4), Murdoch (3), Koussy (2), and others. Tests particular to the critical use of one relation (objective, likeness, spatial, relations *of*

evidence, and *by* evidence (reasoning)), now permit of a finer subdivision. Tests of reasoning, for example, are divisible into

$$V^2 + r^2 + s^2 = 1,$$

where

$$V^2 = G^2 + v_i^2,$$

and where r is peculiar in some way to reasoning. Whether the s factor is regarded as error, or given a meaningful division, will depend upon whether tests are pooled or uniform with respect to form, etc.

Factors such as *ps*, *r*, *k* do not necessarily apply to the relations as such. They may be incidental to them. Thus, Wedek could as easily identify *ps* with temperament; Murdoch's *r* might be connected with the need to "keep a large number of facts in the mind's eye verbally"—many reasoning problems are very easy when we draw or note down the possibilities and do not need to carry the details 'in our head'. Similarly, *k* factor is particular to the 'mental manipulation' of shapes—and not to space-relations as such. Such results are at the mercy of nature—we set out to find one thing, but discover another.

Still further subdivision is possible if test forms are segregated. Pooled analogy tests of psychological ability might be divisible into

$$G^2 + v_i^2 + ps^2 + a^2 + s^2 = 1,$$

where a now refers to the analogy test form in some way. Finally, for a particular form of analogy test, with a particular response procedure, the *nf* factor could be sensibly and meaningfully divided into

$$G^2 + v_i^2 + ps^2 + a^2 + t^2 + e^2 = 1,$$

where t refers to response procedure.

III. FACTOR THEOREMS

To arrive at meaningful factors something of the kind I have described above seems to me to be essential. There would, along such lines, be no disorderly scramble for an 'explanation' of observed fractional factors; the 'explanations' would suggest themselves because of the very nature of the test constructions, although as always, alternative 'explanations' for any factor must be borne in mind. The segregations in Table I are initially purely of classificatory significance: it is the purpose of factor theorems to tell us whether these classifications also have functional, factor, reality. To this end I find that the Spearman *theorems* of general, common, and specific (including error) factors are sufficient for all practical purposes. These theorems are implicit in the fractional com-

ponents considered above. In terms of such theorems the classifications in Table I have been shown to have a certain functional reality. This applies particularly to the *material*, *form* and *response procedure* factors. There is no clear evidence as yet that *relation* factors exist as such; and very little work has been done so far with *sense* factors.

IV. Abilities

Several details require mention, in order to deal reasonably with conditions such as Table I contemplates.

To begin with, everyone will agree, I hope, that factor work of the kind under discussion is concerned with *abilities*. Yet Thomson seems nowhere to define what is meant by an ability, even in his recent *Factorial Analysis of Human Ability* (7). Thurstone, in his *Primary Mental Abilities* (8) asks, on p. 1, what is an ability? But he supplies no answer. Instead he takes it for granted that we all know already. Earlier, in his *The Vectors of Mind* (9), he gave an explicit verbal definition—"an ability is a trait which is defined by what an individual can do". This, he added (p. 48), "implies that there are as many abilities as there are enumerable things that individuals can do. Each ability is, therefore, objectively defined in terms of a specified task." So defined, abilities are altogether too raw. They could include activity like sneezing, or feeling and orexis generally. What is required is the explicit formulation of abilities in terms of factor theorems, test construction and psychological hypotheses. Far from being the raw material for treatment by factor methods, for the discovery of primary abilities, unitary factors and the like, *it is my view that abilities themselves have to be defined in terms of the products of prior factor work.*

The nearest to a formal definition that meets my purpose comes from Alexander (10). He found it necessary to talk about 'functional abilities', which were "the resultants of a number of independent psychological factors, some of which may be common to several abilities, and some of which may only occur in one ability" (p. 117). This is an improvement upon anything hitherto. But there are several weaknesses in it. His supposed 'independent psychological factors' are gratuitous assumptions; and it is not sufficient to classify tests with the same factors as tests of one and the same ability.

To make a beginning with this matter of seeing just what abilities should be in factor analysis, I propose to define them in terms of non-fractional factors.

V. FACTORS

(1) *Non-fractional factors*

We have to distinguish between two kinds of factors and factor methods, fractional and non-fractional. The idea is very simple. Verbal tests require two factors, g and v, to imitate their intercorrelations. These I call fractional factors (f). But the correlations for verbal tests amongst themselves may require only one common factor V to represent them. This is, in a sense, the prior whole, that is divisible into two fractional parts g and v. Whole factors such as V are called non-fractional factors (nf). They are usually correlated with other nf factors, and have no analytical pretensions. Fractional factors, on the other hand, are usually uncorrelated and analytical in purpose.

To the extent that more than *one* factor is involved in representing a score, the factors are necessarily fractional. But one of these may be an error factor (e), or specific (s). I define a non-fractional factor to begin with in terms of two criteria—*if a set of variables satisfy specified psychological conditions, and can be represented by a common factor D and specific factors* (s), *D is by definition a non-fractional factor* (nf). Relative to other variables, however, the nf factor may be divided into additive or other components, and any such will be by definition a fractional factor (f). For convenience f factors are given small letters, and nf factors capitals.

This distinction between nf and f factors may seem very innocuous. Yet the two have very different parts to play in psychometry. The former, for example, define abilities and types for us, whereas the latter do not (5).

Any non-fractional factor, for cognitive activity, will be an ability by definition. nf factors merely represent unanalysed abilities. Fractional factors, unless they are at the same time nf factors, in no way represent abilities—they *describe* them, or might even *explain* them, but do not *represent* them. Just as in Q-technique I seek to draw sharp distinctions between analytical and non-analytical procedures, so also here in r-technique everything is to be gained by refusing to confuse the direct representation of abilities with their fractional analysis into additive components. The methods used for analysis of correlations into nf and f factors are different (see (5), p. 17).

(2) *Relation between nf and f factors*

There is a serious objection to Alexander's 'functional abilities'. Two verbal intelligence tests, for example, may have the following factor loadings:

	g	*v*
Test 1	0·80	0·20
Test 2	0·20	0·80

These would not pass muster either mathematically or psychologically as tests of the same ability. Tests with the same *f* factor composition only satisfy the tetrad criterion amongst themselves when certain relations exist between the factor saturations.

Thus, if four tests $x_1x_2x_3x_4$ have the same fractional factors α and β,

$$r_{x_1x_2}=\alpha_1\alpha_2+\beta_1\beta_2,$$
$$r_{x_1x_3}=\alpha_1\alpha_3+\beta_1\beta_3,$$
$$\ldots\ldots\ldots\ldots\ldots$$

and similarly for $r_{x_1x_4}$, $r_{x_2x_3}$, $r_{x_2x_4}$, $r_{x_3x_4}$.

If we now form the tetrad criterion t_{1234},

$$t_{1234}=r_{x_1x_2}.r_{x_3x_4}-r_{x_1x_3}.r_{x_2x_4}$$
$$=(\alpha_2\beta_3-\alpha_3\beta_2)\,(\alpha_1\beta_4-\alpha_4\beta_1).$$

Similarly,
$$t_{1243}=r_{x_1x_2}.r_{x_3x_4}-r_{x_1x_4}.r_{x_2x_3}$$
$$=(\alpha_2\beta_4-\alpha_4\beta_2)\,(\alpha_1\beta_3-\alpha_3\beta_1).$$

Similarly,
$$t_{1342}=r_{x_1x_3}.r_{x_2x_4}-r_{x_1x_4}.r_{x_2x_3}$$
$$=(\alpha_3\beta_4-\alpha_4\beta_3)\,(\alpha_1\beta_2-\alpha_2\beta_1).$$

These tetrads are therefore equal to 0 when

$$\left.\begin{aligned}
\alpha_2\beta_3-\alpha_3\beta_2&=0,\\
\alpha_1\beta_4-\alpha_4\beta_1&=0,\\
\alpha_2\beta_4-\alpha_4\beta_2&=0,\\
\alpha_1\beta_3-\alpha_3\beta_1&=0,\\
\alpha_3\beta_4-\alpha_4\beta_3&=0,\\
\alpha_1\beta_2-\alpha_2\beta_1&=0.
\end{aligned}\right\} \quad \ldots\ldots(1)^1$$

In practice we could inspect a table of *f* factor saturations, and select from those with the same factor composition those whose saturations are in proportion. When tests involve two or more fractional factors, the

[1] See Burt (11).

proportionality such as (1) expresses is essential for any consistent definition of prior abilities in terms of fractional components.

VI. Conclusion

We can now return to consider Table I. We are faced with an unlimited number of mental tests, actual or possible. Systematic test agglomeration makes for comprehensiveness, and would do far more than anything else to place factor work on a sound footing. I have indicated in Table I how this is pursued in Spearman work, in so far as I have been concerned with this work.

The agglomeration also ensures a certain tangibility to f factor components, and suggests where it is, and where it is not, profitable to try to subdivide specific factors into f factors and error components.

But in practice we have not only to analyse tests; we have to define abilities. A number of tests can have the same factor components but different proportions of loadings. Unless the components are in proportion it is difficult on mathematical no less than on psychological grounds to regard the tests as tests of one and the same ability. One suggestion is that instead of dealing with, for example, verbal tests of many different proportions of factors g and v, we could concern ourselves first with those, say, of two proportions, e.g. (a) tests with g and v loadings in proportion approximately 4 : 1, and (b) others in proportion 1 : 4. These would be defined as tests of two abilities (a) and (b) respectively. This is what I meant by 'defining' abilities in terms of prior factor analysis.

On the other hand, if thorough test agglomeration is pursued the result would be that tests at any particular level of agglomeration would be analysable into components whose saturations would tend to be in proportion. And this I have indicated in Table I by using capital letters for the 'whole' factors at each level of agglomeration: either in fact, or by selection, the original factors in each column of Table I would be non-fractional, and could be defined as abilities.

REFERENCES

(1) Brown, W. & Stephenson, W. (1933). "A test of the theory of two factors." *Brit. J. Psychol.* xxiii, 352.

(2) Koussy, A. A. H. el (1935). "The visual perception of space." *Brit. J. Psychol. Monogr. Suppl.* vii, no. 20.

(3) Murdoch, J. W. (1933). "A factor study of reasoning." (Unpublished Ph.D. thesis, University of London Library.)

(4) Wedek, J. (1933). "A factor study of psychological ability." (Unpublished Ph.D. thesis, University of London Library.)
(5) Stephenson, W. (1939). "Methodological consideration of Jung's typology." *J. ment. Sci.* xcv, 1.
(6) —— (1931). "Tetrad differences for verbal subtests relative to non-verbal subtests." *J. educ. Psychol.* xxii, 334–50.
(7) Thomson, G. H. (1938). *The Factorial Analysis of Human Ability*. University of London Press.
(8) Thurstone, L. L. (1938). "Primary mental abilities." *Psycho-metric Monogr.*, i. Univ. Chicago Press.
(9) —— (1935). *The Vectors of Mind*. Univ. Chicago Press.
(10) Alexander, W. P. (1935). "Intelligence, concrete and abstract." *Brit. J. Psychol. Monogr. Suppl.* no. 19.
(11) Burt, C. (1938). *Psychometrika*, iii, 151.

THE FACTORIAL ANALYSIS OF ABILITY

AGREEMENT AND DISAGREEMENT IN FACTOR ANALYSIS: A SUMMING UP[1]

By GODFREY H. THOMSON

I WOULD like to express my intention to strive for impartiality in this summing up. And with that in view I remind myself first of all that neither Thurstone nor any keen follower of his has been present to defend the idea of 'Simple Structure'. Let me therefore for one paragraph hold a temporary brief on his behalf.

I recently heard Prof. Dirac uphold, before the Royal Society of Edinburgh, the thesis that when a mathematical physicist finds a mathematical solution or theorem which is particularly beautiful—and every mathematician knows what mathematical beauty is—he can have considerable confidence that it will prove to correspond to something real in physical nature. Something of this same faith seems to me to lie behind Prof. Thurstone's trust in the coincidence of 'Simple Structure' in the matrix of factor loadings with psychological significance in the factors thus defined. This is of course not all. The conditions laid down in 'Simple Structure' are laid down in order to remove the embarrassing number of degrees of freedom which permit the factor axes to be rotated like a Catherine wheel. They clamp the axes at one point (unless very unusual conditions prevail in the battery) if indeed that point can be reached, which is not necessarily the case in every battery. If it cannot actually be reached it can be approached as nearly as possible. And Thurstone also gives, in pp. 71 and 72 of *Primary Mental Abilities*, and perhaps elsewhere, what one might call his common-sense reasons for thinking it plausible that numerous zeros will occur among the loadings. Moreover, we must remember that he has embarked on a campaign of forming augmented sub-batteries round tests chosen from his original 57, to obtain better definition of his factors, and has obtained what one must admit to be rather encouraging results in his study of the perceptual factor (*Psychometrika*, March 1938).

The disagreement between Thurstone and Spearman as to the general

[1] A contribution to the Symposium presented at the Extended General Meeting of the British Psychological Society held at Reading in April 1939.

factor *g* is perhaps the most important point of controversy, as it is certainly one of the most easily grasped by the non-mathematical psychologist and one of the most disquieting to him. Prof. Spearman has defended the general factor which will always be associated with his own name; but I think (crossing the court now to speak on his side of the case) that he might have done so more fundamentally than he has. For the objections he has raised to Thurstone's work might conceivably all be overcome. A longer time could be given to the tests. Product-moment instead of tetrachoric correlations could be calculated. Error could be reduced by greater care, and by using large and better samples of subjects. Yet still Thurstone might be able—indeed, with a properly chosen battery, I think he almost certainly would be able—to attain 'Simple Structure' among his loadings, and the complaint that a dismembered *g* was merely being submerged in a sea of error would no longer be justified, at any rate as far as the submergence was concerned. Surely the real defence of *g* is simply that it has proved useful. It still remains to be seen whether Thurstone's primary factors will prove equally useful. A jury in this country would probably to-day give a verdict to Spearman; a jury in another place might give a verdict to Thurstone. The larger jury of the future will, I think, decide by noting which system has proved more useful in the hands of the practising psychologist.

I myself lean at the moment more towards Spearman's *g* and his later group factors than I do to Thurstone's, since they seem to me more in accord with the ideas of my own Sampling Theory. On that theory *g* is as it were the whole mind, and tests are part of *g*, not *g* part of the tests. And were that mind entirely undifferentiated, structureless, *g* would be the only factor we need. As the complexity of the mind, and the complexity of the upper brain, is organized (partly by the maturing of hereditary bonds, mainly I fancy by education and life) and integrated into 'pools', 'clusters', call them what you will, so additional factors, additional descriptive coefficients, are needed. It seems to me at present wise to retain one coefficient to express the general depth starting from which the integration, the deepening into subpools, has gone on. But I am not sure, and think the wiser course is to await further papers from workers in Thurstone's school. I think it possible that there will come a stage when practical progress will be most facilitated by a firm voluntary agreement among psychologists that certain well-known tests have such and such a factorial composition, are not to be changed, and may be used to fix some of the factorial axes in a battery. The sample of persons would also have to be defined and different factorial compositions might be

given for children and adults, for literates and illiterates, etc., etc. I have reason to think that Dr W. P. Alexander would agree with this position.

The debate about the justifiability of negative loadings has been rather inconclusive for two reasons. It has not been fully appreciated, I think, that the objection among certain psychologists to negative loadings is not an objection to one or two such, for a few negatives can be made plausible, as Dr Flugel has shown. It is an objection to having almost exactly half the loadings negative in the factors following the first one, and to the artificiality, which Dr T. V. Moore has pointed out, of explaining even a matrix of exclusively positive correlations by numerous negative loadings. Prof. Burt's defence of negative loadings and of *bipolar* factors is an out-and-out one. He cannot have his reciprocity principle otherwise. His first factor is the average. Reciprocity only comes in among the remaining factors, and it is essential that they be all bipolar, I think, for this to occur. Burt is in fact defending the use of the principal axes of an ellipsoidal space of $n-1$ dimensions, and in any finite battery (whatever may be the case in a battery of *all* tests) will have to meet the criticisms brought against principal axes by Thurstone in *The Vectors of Mind* and elsewhere. There are also certain mathematical peculiarities in Dr Burt's matrix of bipolar marks which must, I think, be further scrutinized. Some of the latent-roots in both Burt's covariance matrices are zero. And when persons and tests are equal in number some really remarkable consequences may ensue. But these are too mathematical to be discussed here.

Moreover, although I agree with a sort of non-rigorous use of the reciprocity principle, I must emphasize that it only can be found rigorously in a very special sample of people who are all average in ability, and a very special sample of tests which are all of average difficulty. Prof. Burt will protest that it will hold approximately elsewhere, and I readily agree that it will do so if these conditions are not too flagrantly broken.

But only if covariances, not correlations, are analysed. And here let me say that I entirely approve of analysing covariances. My criticism has not been directed (and here I think I have Dr Stephenson with me) against the use of covariances as such, but against the arbitrary nature of the units we have been so far compelled to employ. I am particularly gratified to find that Prof. Burt, from his different approach, is so much in agreement with the general principle that I was led to enunciate from considerations arising out of the Sampling Theory, that what I have called the richer (or as he says the more complex) tests should have the larger variance and that what Spearman calls g-saturation is a tentative

measure of their standard deviations if they are in a *g*-battery. I hope we may co-operate in developing this idea.

Dr Stephenson's exposition of what he calls *non-fractional* factors—compounded in fixed proportions of the 'fractional' factors already known to us—is interesting, but too new to be safely judged yet. It would appear, as he says, to have some affinity with the ideas expounded in Dr Tryon's new monograph on 'Cluster Analysis'. But the connexion of these ideas with the main lines of this discussion are not so obvious as to permit of ready discussion here.

Prof. Burt expressed the hope that the 'tetrad difference' of the four symposiasts would be found to vanish! I think we can all agree that some of the 'disturbers' of unanimity have proved not to be 'significant' and that those which remain have been more clearly defined, the first step towards their reduction. We must all be grateful to the organizers of this symposium for these results.

SOME FACTORS IN THE GENESIS OF INTEREST IN PSYCHOLOGY

By R. C. OLDFIELD

From the Cambridge Psychological Laboratory

I. *Introduction* (pp. 109–110).
II. *Interests and interest-tendencies* (pp. 110–113).
III. *Scientific interests* (pp. 114–116).
IV. *Philosophical interests* (pp. 116–117).
V. *Religious interests* (pp. 117–119).
VI. *Divided interests, vacillation, and apparently contingent decision* (pp. 119–123).
References (p. 123).

I. Introduction

"Bekanntlich gibt es Menschen, denen durch eine sei es angeborene, sei es durch äussere Erziehung erweckte Anlage, ihr kunftiger Beruf von früher Jugend an vor Augen steht. Das sind die Glücklichen, die unbekümmert ihre Strasse wandern können, auch wenn sie sonst mancherlei Hindernissen begegnen.... Ich gehöre nicht zu diesen Glücklichen." Thus Wundt ((1), pp. 57–8) and the list of the unlucky includes other psychologists beside him. These others, like Wundt, were not only unable at an early age to experience a sense of vocational certainty, but suffered indecision and change prolonged beyond their natural and customary term.

Conflict in vocational choice may arise in the clash of circumstance with interest. The keenest-minded and the most able may be compelled to turn newspaper boy. William James, born of reasonably well-to-do parents, at one time considered the respective merits of careers in printing and science.[1] But another source of conflict lies within the interests themselves. One interest will urge one kind of activity, direct attention upon one sort of material, induce one variety of cognitive attitude and act. Another demands activity, materials, and cognitive attitudes and acts that contradict or are incompatible with these. Psychologists seem not infrequently to have been the subject of such conflict. Their psychological

[1] See Perry, R. B. ((2), p. 215). The same letter, which is dated 1863, is printed in *Letters of William James* ((3), p. 50).

interests appear often to have arisen in the course of its resolution. And so, perhaps, we may hope to discern, in the division and shift of attention, in the conflict and its resolution, the steps by which a strong interest may develop. We may expect to see the process more clearly displayed here than in the lives of those whom Wundt calls lucky, in whom decision takes place unseen, stirring up no tell-tale agitation of consciousness. To notice instances of divided interest in the earlier lives of those in whom strong psychological interests later developed, and to attempt to relate the reconciliation of these opposing tendencies with the arousal of an interest in psychological problems, will at least provide entertainment for the historically minded.[1] It may in addition supply material for considerations relevant to the general problems of interest-growth.

II. Interests and interest-tendencies

A certain adjustment of terminology[2] will assist our argument. A prevailing tendency in this field would have us define and name psychological functions in terms of the material upon which they are directed. There results a language scarcely adequate for discussion of the dynamic aspects of interest-formation and interaction. For such a procedure assumes that an interest in a specific type of material is a specific psychological function or entity. Relations between interests *qua* psychological functions can, on this basis, be referred to only in terms of relations between the types of material concerned; similarity and difference between interests can derive only from similarities and differences of types of psychological material.

Such language as this, it is fairly plain, is inadequate to the description of the phenomena as we observe them. In the first place, we fre-

[1] As to material, I have relied on (*a*) autobiographies, (*b*) *History of Psychology in Autobiography*(4), and for those individuals, records of whom exist in neither of these forms, mainly on (*c*) Boring's *History of Experimental Psychology*(5).

[2] The remarks that follow are, it will be evident, largely inspired by Prof. Bartlett's views upon interests, and his notion of the 'fitness' of psychological material to psychological functions. For his discussion of these questions *vide Remembering*(6), pp. 113–14, 210–12, 230–3, 236–7, 290–1, 302, 307). There are, however, important respects in which Bartlett's views, and the field of application which he intends to give them, differ from those presented here, particularly in the relation of *function* to *material*; and I must take all responsibility for my own presentation. I should add that as my own remarks are intended to provide a background to the main topic of the paper, rather than a theory of interests in general, dogmatic brevity seemed preferable to exhaustive treatment. In particular, I would not wish ultimately to maintain that interest-tendencies are to be defined solely in terms of types of operation, while, on the other hand, it is evident that the description of material frequently demands reference to operations.

quently wish to express the fact that two individuals may exhibit different kinds of interest in the same type of material. Secondly, we often recognize a degree of persistence and continuity, beneath the surface play of interests, that defies expression in terms of the similarity of material to which attention is successively directed.

No exception can be taken to the use of the term 'interest' to denote a behavioural fact that an individual tends to pay attention to, and to behave[1] in a characteristic manner towards, a certain type of material. But such use can give little warrant for the assumption of psychological tendencies structured and differentiated in accordance with the structure and differentiation of the material towards which the individual so behaves. Accordingly, I shall try to use the word 'interest' only in its purely behavioural sense. For those psychological functions which, operating, produce manifestations of interests, I will use the term 'interest-tendency'.

For the sake of brevity I will characterize 'interest-tendencies' dogmatically in the following terms. They are essentially tendencies *to perform certain kinds of operation* upon psychological material—*to treat material in a certain way*. They are not primarily tendencies to pay attention to any particular kind of material. They produce interest in a particular kind of material because that material is suited to the performance of a particular kind of operation. They may be said to be active in the sense that they reach out for material upon which to operate, choosing that which is suited to them—sometimes, when fitting material lacks, they will mould what stands at hand to fit. They must be supposed to arise early, and to persist unchanged, over long periods in an individual's life. Two or more interest-tendencies may conflict at a given time in the sense either that the types of material upon which they are respectively operating are grossly discordant, implying conflict of a logical or moral nature, or that the activity that each compels is materially or logically incompatible with that compelled by the others. They may accord, or become reconciled, in the sense either that discordant behaviour is not produced, or that the behaviour compelled by each is relegated to separate phases of the individual's life, or, thirdly, that material is found upon which all may operate together.

Accordingly, then, an individual will be supposed to possess constitutional propensities to treat psychological material in a particular way, to perform a particular kind of operation whose nature is in a sense independent of the material upon which it is performed (but see footnote 2 to

[1] The word 'behave' is to be taken to include implicit behaviour.

p. 110). He may possess a number of these propensities various in strength. The environment provides him with a choice of material, and the special circumstances of his life may limit that choice. Propensities, seizing upon material suited to themselves, produce interests. Later, the environmental circumstances may change, the choice of material widen. Material more suited to a particular interest-tendency may offer itself, and a change of interest will result, sometimes dramatic and surprising in the lack of similarity between the two types of material concerned. In many people one interest-tendency is dominant, in others two or more are sufficiently well balanced in strength to induce a conflict of activities. The resolution of such conflict may in certain cases be brought about by the discovery of a type of material suited to the operation of all the interest-tendencies. More generally it may be suggested that the final crystallization of an individual's interests represents the attainment of an optimum equilibrium of the interest-tendencies, in which each obtains the greatest scope possible, having regard to its relative strength, and to the material available.

How then are we to define and name the interest-tendencies? Presumably by the names of the operations, of the ways of treating psychological material, that are characteristic of them. But how are the operations to be defined? In this connexion we must point out that such postulation of determining factors by no means requires that we forthwith define, name and distinguish all the factors postulated. Our procedure will be justified in so far as we are able to characterize, as a first approximation, *any* differentiation within the totality of factors which manifests itself in behavioural consequences. In the second place, the assertion of an underlying differentiation in regard to a limited field of interests, in terms that can apply only to that limited field, carries no implication that the same differentiation fails to hold in regard to other fields, and at other levels. We are concerned here with what, for lack of less invidious terms, may be called the 'higher' or more abstract interests. But it may be the case that, for example, the interest-tendency that determines in part an individual's interest in philosophy may also determine in part his interest in model engineering.

Accordingly I will postulate, for our present purposes, two specific interest-tendencies, the *Empiric* and the *Systemic*. The empiric we may provisionally define as a propensity to treat material in its aspect of relating to experimental fact, the systemic as a propensity to treat material in its aspect of possessing internal structure.

Lastly, we require some means of detecting the operation of these

tendencies. Here we must distinguish between the means that might be adopted in a detailed observational study of the development of an individual, and those we must fall back upon where our material consists in written matter of varied fullness and documentary value. In the former case it would be possible to record not only the material upon which an individual's attention is directed, but also the nature of his activity in regard to that material. It might be possible to see, for example, that while a dominant interest is economics, he is concerned more with the erection of general theories for the explanation of general conditions, than with the determination of empirical fact, suggesting the operation of the systemic rather than the empiric tendency in the arousal and maintenance of his interest.

But here we are driven back upon the argument from the existence of interests to the operation of interest-tendencies. We have seen that to a given interest-tendency some kinds of material are more fitted than others, and that interest-tendencies tend to seize upon material most suited to them. The existence of particular interests, therefore, affords presumption of the operation of particular interest-tendencies, *in circumstances where the range of material at the disposal of the individual is wide enough to allow the factor of 'choice of the fittest' sufficient scope.* Where the range of material is very narrow, as at an early age it may well be, any material, although fundamentally unsuited, may be seized upon and an interest developed. And in such a case the argument from interest to interest-tendency would be highly unreliable.

But in the cases we shall consider this condition is, in general, sufficiently fulfilled. It may be asked why, for the present purposes, we have devoted some pages of discussion to the erection of the concept of interest-tendencies, when admittedly the only means of detecting them in the material under consideration is through the occurrence of interests. We may reply that we shall have occasion to notice the arousal of psychological interests apparently as an outcome of a previous conflict or division of other interests. Now it is difficult to imagine that an interest in psychology can *causally be produced* by the conflict of interests in physics and philosophy. But it is possible to see interests in physics and philosophy as manifestations of two interest-tendencies which, by both seizing upon the material of psychology, produce an interest in psychology, and bring about a resolution of the previous conflict.

III. Scientific interests

Many who have contributed much to psychology must be regarded, in the light of the whole of their achievements, primarily as workers in other fields of science.[1] In such cases there can be little significance for our purpose in tracing the growth of scientific interests which in fact dominated the individual's life. We shall be concerned only to illustrate the fact that *among psychologists proper*, and even among those whose contributions to the subject have been mainly of a philosophical nature, there is a somewhat surprising prevalence of scientific interests.

Bain, up to the time of his entrance into college, was a self-educated man, devoting himself almost exclusively to mathematics and physics, and attaining considerable proficiency. Even while pursuing the multifarious courses required for the Scottish University degree of those days, and at a time when very strong philosophical interests had arisen, he found time to study other scientific subjects, including anatomy (7). Fechner, matriculating as a medical student, turned to physics, and became a professor in this subject, publishing papers of some importance upon the measurement of the E.M.F. of cells. The passage of his interests from physics to psychology was marked in part by publications upon colour sensations and after-images, and in part by a crisis during which he became convinced of a philosophical mission. The desire to set his philosophy upon an empirical basis resulted in the psycho-physical work for which he is remembered ((5), pp. 265–74). Lloyd-Morgan's formal college education was that of a mining engineer ((4), II, pp. 237–64). Drever, in the course of a lengthy education, seems to have studied almost every branch of science, and at one time decided upon a purely scientific career ((4), II, pp. 17–34). James's more academic youthful preoccupations and vocational intentions were biological. He was a member of Agassiz's South-American scientific expedition in the capacity of naturalist. At various junctures he intended to make scientific research and teaching his career. He was medically qualified, and obtained his first academic post as lecturer in physiology at Harvard.[2] McDougall had an extensive scientific education, and confesses that he found difficulty in choosing a branch of science whose problems should match his capacities ((4), I,

[1] E.g. Helmholtz, Galton, Hering, Loeb. But I have not, in selecting material, had too strict regard to a proper definition of the term 'psychologist'.

[2] James's life constitutes perhaps the most striking illustration of the points raised in this paper. Besides the *Letters* (3), we now have the magnificent source provided by Perry (2) in which a mass of previously unpublished letters are nicely blended with wise narrative.

pp. 191–223). C. S. Myers, whose Cambridge education was scientific and medical, was at one time nearly drawn to pathological research as a vocation ((4), III, pp. 215–30). Ward, the purest of the pure in his later writings, sowed a few youthful wild oats in natural science (8), publishing on the nervous system of the crayfish, and upon the vitality of the common snail (9). Lotze's scientific interests were sufficient to allow him to select his father's profession of medicine, although even during the time of medical training he was "...by temperament...more disposed towards the arts and philosophy than towards science and medical practice" ((5), p. 251).

Even among those who never seem to have contemplated a career in one of the well-established natural sciences, there are a number who went out of their way in student and early post-graduate days to study science. Fröbes, while immersed in the exacting educational programme of the Society of Jesus, a programme in which dogmatic theology and its contributory philosophy predominate, pursued the physical sciences with enthusiasm ((4), III, pp. 121–52). Titchener, having read "Greats" at Oxford, promptly secured a research studentship in physiology, later writing on biological topics.[1] Stanley Hall, even as late as his second visit to Germany, mixed in a fair proportion of science among the numerous courses on almost every kind of subject that he attended, and found time for experimental work in neurology with von Kries ((12), (13), p. vii).

A factor which quickened the scientific interests of philosophical students round about the second half of the last century, particularly in Germany, was a prevailing tendency to regard science as essentially the true method of philosophy (see (5), p. 352). In consequence we find Stumpf, instigated by Brentano, undertaking "...to learn science in the chemical laboratory, even to the extent of setting the building on fire..." ((5), p. 352), while Woodworth, less disastrously (at least so far as his personal confessions indicate), pursued a similar course with the same motive, later hesitating between psychology and mathematics as a career, and doing a year's research in the physiological laboratory at Harvard, before settling down as a psychologist ((4), II, pp. 359–80). The experimental work for which Fechner and Ebbinghaus are remembered originated in part in the same tendency.

Strong interest in mathematics appears to be a less common occurrence. In only one or two cases is interest so strong, or ability so great,

[1] One may perhaps be permitted to express the sincere hope that Prof. Boring's next achievement will be a *Life* of Titchener. For meagre, but appetizing accounts, *vide* Boring ((5), pp. 402–13); Boring(10) and Warren, H. C. (11).

that the individual considered finding an academic vocation in this subject. Woodworth and Yerkes are exceptions.

It is clear, then, that scientific preoccupations played an important, and sometimes a dominant part in the development of interests in a number of those who finally turned to psychology. Frequently these preoccupations induced the individual to decide upon a career in science, or caused him to undertake scientific research. And they are to be found even among those, like James, Lotze and Ward, whose later work tended more and more towards the purely philosophical.

In a number of the instances we have noticed, it would seem reasonable to interpret the existence of scientific interests often co-existing with philosophical as the manifestation of an empiric interest-tendency. For it is often possible, without straining the argument too far, to see the later operation of the same tendency even where pursuits other than the purely scientific have become dominant. James's philosophy is shot through and through with evidence that, even in the field he finally chose, a strong empiric tendency found the means of expression. Fechner, in the moment when a fervent, almost mystical, sense of having a philosophical mission seized him, sought an empirical basis for his system. Titchener's empiricism underlies his structuralism and his passionate rejection of impalpable cause-factors. It would scarcely be too much to say that Lloyd-Morgan's Canon was forged from ore of his own mining.

IV. Philosophical interests

It is not in itself a matter for surprise that, with very few exceptions, the biographical material makes mention of more or less strong philosophical interests at one period or another in the lives of the individuals. Those about whom information is available belong to older generations. In the times of their youth psychology was approached through, and regarded by many as, a branch of philosophy. Again, while it is in the very nature of natural science, by one device or another, to eliminate psychological issues from its content, philosophy has been less successful in this respect, and the pursuit of metaphysics suggests problems of psychology at every turn. Accordingly, the instances of philosophical preoccupations which may be considered significant from our point of view, are those in which they accompany other, generally scientific, activities, and other educational programmes and vocational intentions.

Binet was strongly attracted towards philosophy while pursuing the study of medicine(14). Janet, effecting a compromise between strong scientific and theological tendencies, turned to philosophy and became a

lycée professor in that subject at the age of 22 ((4), I, pp. 123–33). Lloyd-Morgan, while applying himself to the vocational training of a mining engineer, "... kept close, though necessarily occasional, touch with Berkeley and Co...." ((4), II, p. 240). G. E. Müller studied history because he intended to teach it as a profession, and philosophy because he liked it ((5), p. 361). C. S. Myers acquired philosophical interests while an undergraduate studying science and medicine ((4), III, p. 216), and later published papers on purely philosophical topics. Spearman confesses to "... an excessive but secret devotion to philosophy..." before and during his school days, a devotion which survived an army career ((4), I, pp. 299–333). Lotze's philosophical interests are said to have arisen at an early age, and he continued to pursue them throughout a brief medical career ((5), pp. 250–4). James's numerous preoccupations as a young man included philosophy. The Agassiz Expedition interrupted his scientific and medical studies at Harvard, and in letters from Brazil, "behind his more or less playful manner there is unmistakeable evidence of a brooding preoccupation with philosophy" ((2), p. 215). He "...contributed little to the work of the expedition, but he made in Brazil one important discovery: he found that he was a philosopher" ((5), p. 495).

An interest in philosophy, then, is not uncommon even among those psychologists whose original vocation ranged from soldiering to medicine and mining. And it is reasonable to interpret the maintenance of such interests, concurrently with other pursuits, as evidence of the operation of some general systemic interest-tendency.

V. Religious interests

The fact that mention is made of religious factors in a large number of cases where biographical material is available is not surprising.[1] In many instances the reference relates to a religious family or community background, and the same might be found in any other set of comparable records of intellectual individuals in the latter half of the nineteenth century. It is interesting to notice, however, that some individuals attribute their psychological interests to earlier religious preoccupations.

[1] Mention occurs in about half the total cases. Thoughts of entering the ministry seem to have been entertained by only ten of the individuals concerned, and of these three only reached the stage of ordination. Two of these subsequently seceded from the Church, the remaining one being Fröbes. Brentano's ecclesiastical career was attended with results more stimulating to the Church than satisfactory to himself, and Ward's incumbency of a Nonconformist Chapel was brief. In three cases, McDougall, Brentano and Stumpf, there is definite mention of a religious crisis.

Höffding states that he was led "...partly through the problems of religious emotions to speculate on the mind of man" ((4), II, p. 197). Judd writes: "I have no doubt that my early intensive religious training had a large influence in directing my later interests towards a study of man and his mental life" ((4), II, p. 207). Janet is of the opinion that his interest in psychology and philosophy was the result of a conflict between two tendencies, the one religious and quasi-mystical, the other scientific ((4), I, p. 123). Baldwin writes that his "...interest was directed towards psychology through his early intention to go into the ministry" ((4), I, p. 1). Freud declares: "My early familiarity with the Bible story (at a time almost before I had learnt the art of reading) had, as I recognized much later, an enduring effect upon the direction of my interest" ((15), p. 13).

What part can we suppose religious preoccupations to play in the later development of psychological interests? There can be no denial of the fact that the content of religious knowledge and experience may in itself suggest problems of psychological import. But it may be that this fact does not exhaust the possible connexions between religion and psychology. Religious teaching frequently provides for the growing child the first body of more or less coherent psychological material. It offers admirable opportunity for the working of the systemic interest-tendency, as any casual study of children's theological systems will show. I will suggest that early religious interests afford evidence of the operation of the systemic interest-tendency, later to find a nicer balance with the empiric in becoming directed upon the problems of 'man and his mental life'.[1] In this suggestion there is, of course, no implication that all early

[1] A problem perhaps relevant is that of the significance of experiences and phantasies, apparently not uncommon in intelligent children, incorporating religious, mystical, and astronomical or cosmological elements. The experience may consist merely in a tendency to preoccupation concerning the ultimate significance of the universe. Another variety is described vividly by Klemm ((4), III, p. 155) and is worth quoting. "The first connected realm of facts that presented itself to me was the kingdom of the stars. The boy's fantastic thought, kindled no less by the infinity of space and time than by the unattainable and supramundane, rose upward through the silent winter nights. The contradictions of religious dogma were challenged and experienced over and over again with a peculiarly sweet tension. When I was 15 I received a gift of a three-inch spyglass with a lens of ×80 magnifying power, and with the help of wood and paper constructed a telescope." The combination of an almost mystical experience with a truly empirical curiosity, seems to be characteristic. Another instance is provided by the highly systematic fantasies of voyages through the solar system developed by Wirth and his brother(16). To pursue the matter further here is impossible, but one is tempted to suggest that such phantasies may represent the operation of active interest-tendencies upon such material as lies at hand, material which is not truly fitting.

religious interests are the product of the systemic tendency, or that it alone can ever be effective in their production.

VI. Divided interests, vacillation, and apparently contingent decision

We come now to an aspect of the subject which can scarcely fail to strike forcibly anyone who studies the biographical material, however superficially. It reveals in a surprising degree the prevalence of conflict of interests among psychologists. Such conflicts are manifest in a division of strong interests often persisting beyond the phase of individual development in which they are ordinarily considered normal. They are shown most clearly in a tendency to vacillate between the adoption of different intellectual pursuits, and in some cases are indicated, one may suspect, in the apparently contingent nature of the particular event that leads to a decision.

Division of strong interest, and a tendency to vacillate between different intellectual pursuits, are apparent in many cases. Külpe intended to study history in the first place, but going to Leipzig for this purpose, he came into contact with Wundt, and was for a time diverted to experimental psychology. He then went to Berlin where he read history with Mommsen. This phase lasted for a single semester, at the end of which he returned to psychology, this time with G. E. Müller at Göttingen. However, still undecided as to whether History or Psychology should be his life-work, he repaired to Dorpat, where he again read history. In 1887 he took his degree in psychology, finally deciding in favour of this subject ((5), pp. 386–7). In his later years he turned to philosophy.

G. E. Müller's primary interest at the age of 15 was philosophy, and at this time, he tells us, he was inclined to mysticism from which he was rescued by the writings of Lessing. As a student he read history for the sake of a teaching career, and philosophy because of a strong interest. After a period of history, archaeology and the History of Art, perusal of Lotze's writings caused him to study philosophy as *Hauptfach*, and a considerable conflict between historical and philosophical interests ensued, aggravated by the belief of the times that the scientific method, in which he had had no training, is the true method of philosophy. The conflict seems to have been painful, but was resolved by the outbreak of the Franco-Prussian war. After a year of service in the field Müller found history a 'narrow' subject, and turned to the study of science ((5), pp. 361–2).

Wundt's decision to specialize in the study of physiology was the

result of painful reflections upon the manner in which he had dissipated his efforts during his first student year, coupled with the courage conferred by the beauty of a boat trip up the Neckar ((1), pp. 71–2). Happily for experimental psychology this decision did not too rigidly determine his subsequent interests.

Such instances of conflict and vacillation are interesting in themselves. But more significant is the question of the manner in which they issue, and are resolved. The resolution, when it comes at all in definite fashion, may happen by one of three more or less distinct means. In some cases an event apparently contingent, and irrelevant to the real issue, may seem to bring about a decision. In others, where there are two or more conflicting interests, it may prove possible in the ultimate event for each to be separately pursued. Lastly there may come a solution through the discovery that a new interest—in the cases we are considering generally an interest in psychology—offers a compromise and reconciliation, allowing the tendencies harmonious expression in unified activity upon a single field of material.

Personal influence may operate in conjunction with an event apparently so contingent and irrelevant that one can scarcely help inferring a pre-existing instability of interest.

Carr, having decided to major in mathematics, arrived at Colorado University, and found that he disliked at sight the instructor in his chosen subject. Happening to meet the professor of psychology however, and finding him genial and encouraging, he decided on the spot to turn psychologist, "...although I knew practically nothing about the nature of the subject" ((4), III, pp. 69–72). Seashore, whose main interests during his early college days were mathematics and languages, was inspired by friends to go to Yale in order to study philosophy. He happened, however, to arrive there on the day the psychological laboratory was opened, and, on the advice of Ladd, he entered the course, although he too knew nothing of it previous to his arrival ((4), I, pp. 225–97). Stern, intending to study philosophy and philology, soon realized that "...the petty pursuits of philology could not thrill me". A deep impression was made upon him by Ebbinghaus, whose point of view appealed to Stern's love of the empirical. Stern had, he says, "...a strong desire for exact and detailed work, and direct contact with the concreteness of fact". Philosophical speculation, however, continued to be an interest, although he regarded it as a weakness ((4), I, pp. 335–88). Dodge started life as a philosopher, but in Erdmann's seminar "...the coincidence of a...humiliating experience and an unprecedented piece of

good fortune diverted me into experimental psychology". The first of these events consisted in Dodge missing a rather obvious point in connexion with Kant, upon which Erdmann remarked: "Herr Dodge, I fear you will never make a philosopher". The second was Dodge's success in inventing for Erdmann a tachistoscope of elaborate requirements, a technical triumph which Erdmann had been assured by his physicist friends was impossible ((4), I, pp. 99–121).

Solution through the continued but separate pursuit of the conflicting interests is not for our purpose of any great psychological interest. It is sufficient to note that circumstances may allow of it, the second interest-tendency obtaining outlet in hobbies and spare-time occupations. But the type of solution in which incompatible tendencies find reconciliation by becoming directed upon some new topic appears commonly as the means by which individuals are led to the pursuit of psychology.

Janet speaks of the general direction of his philosophical and psychological interests as being determined by a sort of conflict. One set of tendencies towards natural science, botany in particular, amounted to a passion which "...determined my taste for dissection, precise observation and classification". Another tendency, never wholly satisfied, was religious, and Janet confesses that his mystical impulses were never wholly controlled. Between the two opposing forces an interest in philosophy arose as a compromise ((4), I, p. 123). It may be supposed that the empiric tendency found insufficient outlet in philosophy, and determined the later change in Janet's interests towards psychopathology.

Washburn writes that "At the end of my senior year, I had two dominant intellectual interests, science and philosophy. They seemed to be combined in what I had heard of the wonderful new science of experimental psychology" ((4), II, p. 338). Woodworth, following a period in which he had devoted himself variously to philosophy, mathematics, and science, remarks: "And so...I turned from philosophy, half expecting to continue some effort in it, but discovering as time went on that psychology was an amply sufficient occupation" ((4), II, p. 367).

Fechner, medical student turned professor of physics and humanistic satirist, suffered a crisis of health which left him with a "...deepened religious consciousness" and a philosophical mission. He felt at the same time that philosophy required a scientific basis, and in 1850 "He saw that the thing to be done was to make the relative increase of bodily energy the measure of the increase of the corresponding mental intensity...". This conclusion led to the experimental work forming the basis of the *Elemente der Psychophysik*, which will be remembered when the satires of

'Dr Mises', and the sevenfold call of the philosopher Fechner to a dormant public, are forgotten ((5), pp. 267–70).

Klemm writes that at an early age "...the joy which I had in scientific insight was perturbed by the philosophical deductions especially of the materialistic school of thinkers". At Munich, where he had enrolled in Lipps's psychology course, "One thing struck my attention, the small collection of apparatus for experimental psychology, which stood unused and somewhat dust-covered, in a corner of the seminar room....Here I felt the presence of a gap, through this gap I passed to another great field, that of experimental psychology ((4), III, pp. 153–61).

Stern's philosophical and empirical tendencies found a compromise in psychology. At the age of 19 he wrote in his journal: "In Philosophy I must find my salvation or perish. One consolation there is, namely, that a philosophical discipline is open to me as my special subject, psychology..." ((4), I, p. 340). Stumpf, after a period of development which included strong and persistent interests in music and philosophy, and a religious crisis of the most intense and agonizing kind, finally decided to devote himself "...to that field which, connecting my musical experiences with the interests of psychology, seemed to me, personally, the most promising". In his case, however, there is evidence of a continued clash of empiric and systemic tendencies, for he mentions that, with Aristotle, he finds theory the sweetest of all, but it is always a comfort to pass to observation ((4), I, pp. 396–7).

James's conflicts of interests and difficulties of vocational decision, aggravated by a superbly delicate moral sensibility, are too numerous and complex for brief record to do them justice. One quotation must suffice to illustrate the intensity of the doubts to which he was subject. He writes in 1868, at the age of 26: "I have not got started properly on any line of work yet, but am hovering and dipping about the portals of psychology. The fact is that I am about as little fitted by nature to be a worker in science of any sort as anyone can be, and yet...my only ideal of life is a scientific life." But later in the same letter he decides that "All in all, even the sweepings of morality are better than chemical reactions", and adds, "Thus do I lash my tail and start myself up again..." ((2), I, pp. 286–8).

Many factors may contribute towards the growth of an interest in psychology. It is an interest which, for reasons relating simply to its present academic status, is in any case liable to develop later than are interests, for instance, in philosophy, history or biology. But one means by which it may be reached is, I suggest, through the existence and

balance of two generalized tendencies, the one to treat material in an empiric, the other to treat material in a systemic, way. In such a case there may occur a competition of activities and a conflict of interests. The situation then becomes resolved through joint direction of both tendencies upon the material of psychology, which can offer scope for the two together.

REFERENCES

(1) WUNDT, W. (1920). *Erlebtes und Erkanntes.* Stuttgart: Kröner Verlag.
(2) PERRY, R. B. (1935). *The Thought and Character of William James,* vols. I and II. Oxford University Press.
(3) JAMES, H. (Ed.) (1920). *Letters of William James,* vols. I and II. London: Longmans.
(4) MURCHISON, C. (1930, 1932, 1936). *History of Psychology in Autobiography,* vols. I, II and III. Clark University Press.
(5) BORING, E. G. (1929). *A History of Experimental Psychology.* New York, Century.
(6) BARTLETT, F. C. (1932). *Remembering.* Cambridge University Press.
(7) BAIN, A. (1904). *Autobiography.* London: Longmans.
(8) BARTLETT, F. C. (1937). "Cambridge, England: 1887–1937." *Amer. J. Psychol.* L, 97.
(9) TITCHENER, E. B. & FOSTER, W. S. (1912). "A list of the writings of James Ward." *Amer. J. Psychol.* XXIII, 457.
(10) BORING, E. G. (1927). *Amer. J. Psychol.* XXXVIII, 489.
(11) WARREN, H. C. (1927). *Science,* LXVI, 208.
(12) HALL, STANLEY (1923). *Life and Confessions of a Psychologist.* New York: Appleton.
(13) —— (1912). *Founders of Modern Psychology.* New York: Appleton.
(14) SIMON, T. (1912). *L'Année Psychol.* XVIII, 1.
(15) FREUD, S. (1935). *An Autobiographical Study.* London: Hogarth.
(16) WIRTH, W. (1931). "Wie Ich zur Philosophie und Psychologie kam." *Arch. ges. Psychol.* LXXX, 452.

(*Manuscript received* 15 *January* 1939)

THE EFFECT OF ADAPTATION ON THE DIFFERENTIAL THRESHOLD FOR SOUND INTENSITY

By A. F. RAWDON-SMITH[1] and R. S. STURDY

From the Psychological Laboratory, Cambridge

I. *Introduction* (pp. 124–126).
II. *Apparatus* (pp. 126–129).
III. *Procedure* (pp. 129–131).
IV. *Results* (pp. 131–135).
V. *Conclusions* (pp. 135–137).
VI. *Summary* (pp. 137–138).
References (p. 138).

I. Introduction

The influence of intense stimulation on the absolute sensitivity of the human ear has been the subject of several previous reports (Rawdon-Smith, 1934, 1935, 1936). In these, it was shown that pure tone stimulation at relatively high loudness levels yielded a transient loss of absolute sensitivity frequently of considerable extent, and exhibiting several interesting features. Thus the extreme lability of the phenomenon and its partially binaural nature (by which is meant the tendency for stimulation of one ear alone to modify the sensitivity of both) led to the supposition that the effect was partially mediated by changes in the central nervous system. Indications of a further peripheral change have, however, also been shown both for the human subject (Rawdon-Smith, 1934) and for the physiological preparation (Rawdon-Smith, 1938, p. 117). The former hypothesis, namely, that of the central locus of certain of the observed phenomena, achieved confirmation in the finding that under certain circumstances it is possible for an unexpected stimulus, even one in an alternative sensory mode, to cause a sharp, though temporary, recovery of normal sensitivity (Rawdon-Smith, 1935, 1936). Since this effect clearly resembles the general phenomenon of 'disinhibition', we have preferred to employ the term 'auditory inhibition' rather than the more conventional 'auditory fatigue', to describe the loss of sensitivity produced by the initial stimulation.

[1] Senior Student for the Royal Commissioners of the Exhibition of 1851.

In view of these several unexpected properties, it has appeared desirable that our investigation be extended to determine whether or not intense stimuli of the type found to yield so great an effect on the absolute sensitivity might not to some extent influence the differential sensitivity of the human auditory mechanism. The present account, therefore, deals with the results of an investigation of the effect of such stimuli on the differential intensity sensitivity of the ear.

Earlier work on the differential intensity threshold for the human ear includes that of Knudsen (1923) and Riesz (1928). As a source of sound, Knudsen used a thermionic tube audio-oscillator, the output of which passed via a galvanometer and resistance network to a telephone installed in a sound-proof room. A motor driven key-switch operated to short-circuit one of the resistances of the network, and at this point the intensity of the sound in the telephone was increased. The amount of this intensity rise could be altered by adjustment of the value of the short-circuited resistance. Starting with an easily supraliminal change, this value was gradually decreased until the just noticeable difference was found. Unfortunately, however, the introduction of an instantaneous intensity rise in the manner discussed carries with it certain important limitations. When either the absolute intensity or the intensity change or both are large, the arithmetical change of force on the diaphragm at the moment of intensity change is sufficient to produce an audible note of the natural frequency of the diaphragm of the telephone in use. This phenomenon is objectionable, as it produces a readily audible click, when the intensity change is effected. These theoretical suppositions were noted by Ewing & Littler (1935) in a related investigation; and were adequately confirmed during certain work preparatory to an investigation of the variation of the value of the differential intensity threshold, with rate of change of intensity (Rawdon-Smith & Grindley, 1935). If Knudsen's technique is adopted, therefore, as the value of the intensity change is reduced so will the loudness of the click be reduced. A human subject of such an experiment must, therefore, make any judgment as to the perceptibility of the change in question at least partly on the basis of the loudness of the click accompanying the change.

Knudsen's work was repeated by Riesz (1928), using a technique against which these criticisms cannot be levelled. This work is discussed in detail by Fletcher (1929, pp. 147–51) and need not, therefore, be further mentioned here, save to note that the very considerable technical elaboration which Riesz was able to introduce would ordinarily render investigations of this character somewhat beyond the scope of most

laboratories. We have, therefore, evolved a simpler technical arrangement, carrying with it the advantages of Riesz', and additionally conferring the benefit of giving a pure intensity modulation, constant in regard to frequency.

II. Apparatus

A sinusoidal alternating voltage, generated by a beat frequency oscillator, is taken through a calibrated attenuator to a special amplifier. The output from this, fed to a telephone receiver, produces an approximately pure tone. The block diagram (Fig. 1) shows the layout of this equipment.

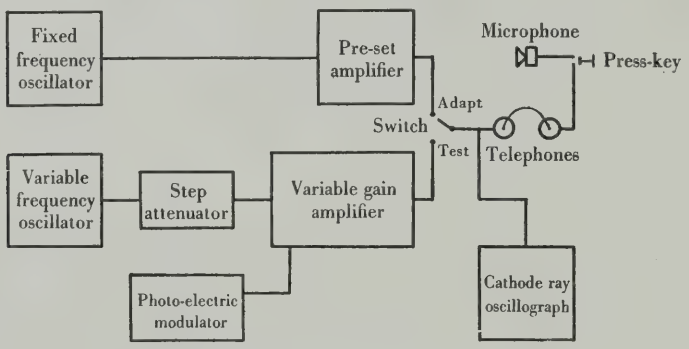


Fig. 1. Block diagram showing layout of apparatus.

The amplifier contains one stage whose amplification can be rhythmically varied by suitable application of a voltage derived from a photoelectrical modulator. It will readily be seen that a rhythmic change in amplification of known amount will produce a tone whose intensity is modulated over a predetermined range.

The design of the modulator is such that no click, or other spurious indication that the loudness of the tone has been altered, is audible in the telephone receiver. It consists essentially of a disk of special shape, rotated in the path of light rays falling on a photoelectric cell in such a way that the amount of light falling on the cell (and consequently its electrical output) is varied in a rhythmic manner. The output voltage so generated is directly coupled in such phase as to decrease the standing negative bias on the suppressor grids of a pair of Mazda ACSP 1 pentodes, operating in biphase as part of the amplifying chain (Fig. 2). The stage gain is consequently increased when more light falls on the photoelectric cell.

For the purpose of the present experiment, a sinusoidal intensity modulation of the tone in the telephone receiver was deemed to be desirable, because such an intensity change scatters least energy into parts of the spectrum remote from the test frequency. This necessitated a sinusoidal change of voltage from the cell, since the amplification of the ACSP 1 valve bears a very nearly linear relation to the applied voltage on the suppressor grid. To obtain this sinusoidal voltage, a cardioid disk (Fig. 3*a*) must be rotated in the path of parallel light rays

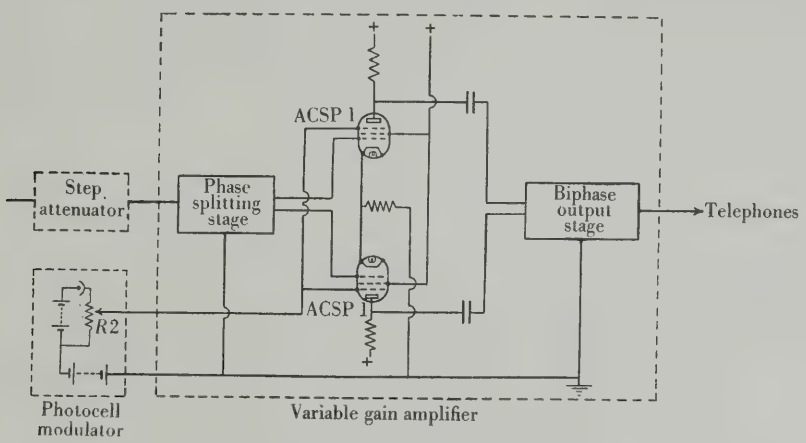


Fig. 2. Circuit diagram showing connexions of photocell modulator and variable gain amplifying stage.

falling on the photoelectric cell. The arrangement employed (Fig. 3*b*) consists of an optical bench carrying a light source, a condenser, a slit to cut off stray light, and the photoelectric cell itself. These can be moved relatively to one another and to the rotating disk, thus enabling correct adjustment to be found very simply. The disk is carried on a rod, and is rotated by a synchronous motor through a belt reduction drive arranged to give one complete rotation of the disk per second; each full intensity modulation cycle thus lasts 1 sec., this modulation frequency having been found by preliminary experiment to be most suited to the observers.

The depth of the modulation is varied by altering the intensity of the light source, and a variable resistance (*R* 1, Fig. 3*b*) in series with the battery supplying the bulb is used for this purpose. The proportion of the output from the cell applied to the grids of the ACSP 1 valves is adjusted on *R* 2 (Fig. 2) until an intensity change subliminal to all subjects occurs with the dimmest light used, and a suitable maximum intensity modulation depth is attained with the full brightness. Varying

the intensity of the light thus gives a smooth change in the depth of modulation, and it was this control which was used in taking readings, in the present series of experiments.

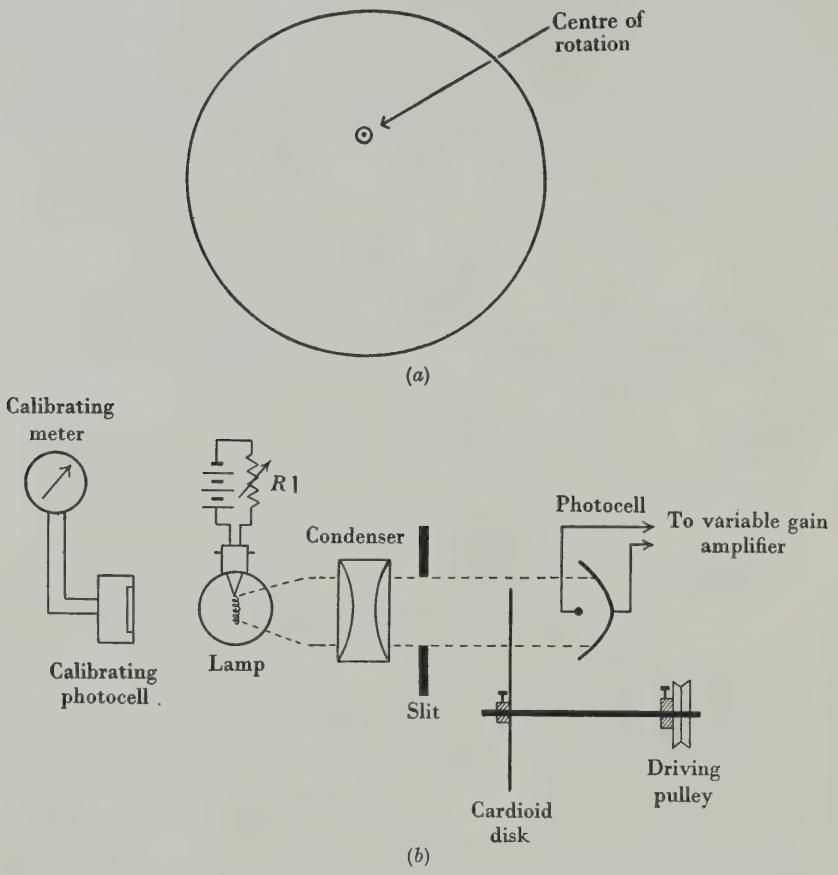


Fig. 3. (*a*) Shape of cardioid disk. (*b*) Arrangement of light modulator: light from the lamp passes through the condenser and slit, and falls on the photocell; the parallel rays are rhythmically interrupted by the rotating cardioid disk. A second, generator type, photocell, and a microammeter, are used to measure the intensity of the light, which is varied by *R* 1.

With the light source at its dimmest, the observer reports no change in the intensity of the sound in the telephones; slowly increasing the intensity of the light increases the depth of modulation, and a point is reached at which the observer reports a rhythmic change in the loudness of the tone. At this point, a reading proportional to light intensity is

taken by means of an independent generator type cell and microammeter (Fig. 3*b*) since a record of this is essential for the calibration of the apparatus.

Calibration is accomplished as follows: a number of readings is taken, and the average light intensity at the threshold point determined. The light source is then set at this value, and a rectifier type output voltmeter switched into circuit across the telephone receiver. The output voltage is read with the cardioid disk rotated first so as to give minimum output, and secondly, so as to give maximum output. The ratio of maximum to minimum voltage across the telephones at the differential threshold has thus been determined, and this ratio may be expressed as follows:

$$\text{voltage ratio} = V + \Delta V : V \qquad \text{.....(i)}$$

(where V is the voltage output at minimum, and ΔV is the voltage output at maximum).

The method of calibration outlined above has the advantage that variations in the voltage supply to the light source, ageing of the bulb, etc., are of no consequence, since it is the actual brightness which is noted at each reading, by means of the auxiliary photocell equipment; moreover, since the output voltage is measured immediately after taking readings, long-period changes in the amplification constants are also unimportant.

The output to the telephones is also applied to a cathode ray oscillograph, and during the course of an experiment, careful check is kept on the waveform of the intensity modulation, and on the frequency of the tones employed. The latter is checked by obtaining a Lissajous's figure against a constant frequency source applied to the other deflectors of the cathode-ray tube.

A voltage from a second, fixed frequency, oscillator (Fig. 1) amplified by a pre-set amplifier whose output stage is similar to that of the main amplifier described above, could be switched into the telephone receiver at will, in place of the output from the main amplifier, and the tone so produced was used as the adaptation tone in the present experiments.

III. Procedure

The observer was installed in a sound-proof room, inside which the external noise level was below threshold. He was seated at a table, on which was a silent, mercury-operated key, pressure upon which sounded a buzzer in the experimenter's room. An additional microphone and

loud-speaker system enabled comments by the observer to be heard by the experimenter. The observer wore a pair of telephones on an adjustable head band, in one of which he heard the experimental tone, and in the other of which he could hear the experimenter's voice, provided the latter pressed a key in order to speak to him. The telephones were marked, so that the experimental tone was always sounded in the same telephone on the same ear throughout any one series of experiments.

The absolute threshold of hearing was first assessed by ordinary audiometric methods, using the system of attenuators in association with the main amplifier, at the frequency at which the subsequent tests were to be carried out. This determination was essential in order to ensure that the differential judgements could be carried out at an intensity level referable to the observer's own absolute threshold.

This preliminary measurement was accomplished in 2–3 min., and the observer was then warned that he would receive an adapting tone for 2 min., after which he would be switched noiselessly to the test tone, whose intensity modulation, inaudible at first, would gradually be increased until he signalled that he was able to detect the rhythmic intensity change. Immediately a judgement had been given, the adaptation tone was switched on for a further 15 sec., and then another test taken; this procedure was continued until five readings had been taken, to enable a fair average to be made, after which (the adapting tone still being kept constant) a further five readings were taken at each of a number of different test intensities. This occupied one complete session. It was thus possible from the data of a single session to plot the differential threshold at each of a number of test intensity levels ranging from 10 to 110 db. above the absolute threshold, when the ear was adapted to a certain adaptation intensity level.

In all, seven sessions were required of each observer, each one being undertaken at a different adaptation intensity, so that the differential threshold, at stated test intensity levels from 10 to 110 db. above the absolute threshold, was known under conditions of adaptation intensity varying from silence to 110 db. above the absolute threshold.

During the course of a session, the observer was allowed to make any comments he wished and these were heard by the experimenter through the microphone communication channel. It was thus possible for the experimenter to disregard any readings which were acknowledged by the observer at the time to have been unreliable. In a session lasting 20 min. or more, it was almost inevitable that the observer's attention would wander, or that sounds of his movements in breathing would on

occasions mask the test tone (more particularly at low intensities). When such events occurred, he communicated with the experimenter, who was thus able to make an appropriate note.

Three observers only were used in the present experiments, owing to the length of time taken in obtaining a full series for each, and to the fact that inexperienced observers, some of whom were tried in an incomplete series, had difficulty in presenting consistent judgements. All three observers employed gave absolute thresholds within 4 db. of each other, and closely corresponding to the normal value for the average human ear at the frequency used.

IV. Results

The voltage ratio of maximum to minimum of the intensity modulation at threshold, obtained as explained above (p. 129), may be treated in either of two ways, as follows:

(*a*) The actual increase in sound intensity necessary to elicit a judgement, expressed in decibels, is simply obtained from formula (ii) below:

$$\text{Differential threshold (decibels)} = 20 \log \frac{V+\Delta V}{V}. \quad \text{.....(ii)}$$

(*b*) Weber's fraction, $\Delta I/I$, can be obtained from the data in the following way:

$$\text{Voltage ratio} = \frac{V+\Delta V}{V}, \quad \text{.....(i), (p. 129)}$$

therefore $$\text{Energy ratio} \quad \left(\frac{I+\Delta I}{I}\right) = \frac{(V+\Delta V)^2}{V^2},$$

hence $$\text{Weber's fraction} \quad \frac{\Delta I}{I} = \frac{(V+\Delta V)^2 - V^2}{V^2}$$

$$= \frac{2V\Delta V + \Delta V^2}{V^2}. \quad \text{.....(iii)}$$

It is convenient to use Weber's fraction as a measure of the differential threshold, and if these values are plotted as the ordinates of a graph (on a logarithmic scale) the abscissae being the intensity levels of the test tone in decibels above the absolute threshold, graphs may be drawn showing the differential threshold at various test intensities for a given intensity of adaptation tone.

A typical graph (Fig. 4) of this type, plotted from the results of a single session by one of the observers for an adapting tone at an intensity level of 70 db. above the absolute threshold, shows the general char-

acteristic of the differential threshold. It will be seen that the curve corresponds to the generally accepted shape of the Weber's law curve when $\Delta I/I$ is plotted against I for any sensory mode, but it is of interest to note that intensity discrimination is finest when the test is made at the same intensity level as the adaptation. In general, this characteristic may be observed for all values of adapting intensity, and if the curves for a number of adapting intensities are plotted together on the same graph (e.g. Fig. 5 A, B or C), it can be seen that each curve tends to approach its minimum value of $\Delta I/I$ when test and adapt intensities are equal.

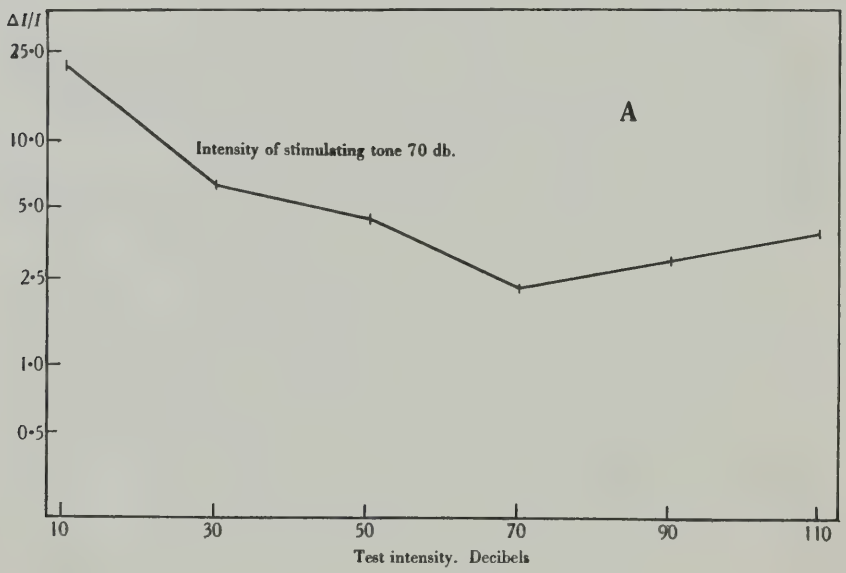


Fig. 4. Curve relating $\Delta I/I$ to test intensity, at 70 db. adaptation intensity. Frequency 1000 c.p.s.

The graphs obtained for each of the three observers are given in Fig. 5 A, B and C. It can be seen that the minimum value of $\Delta I/I$ at equal test and adapt intensities appears on nearly all the curves, but there are some exceptions, notably the 100 db.[1] adaptation curve for A (Fig. 5), which reaches a minimum at 70 db. test intensity. These exceptions would all disappear if one reading were altered: in the present

[1] It must be understood that where this term occurs without qualification in connexion with intensity levels, it refers to decibels above the observer's own absolute threshold at the frequency employed: the decibel expresses a ratio, and not an absolute value.

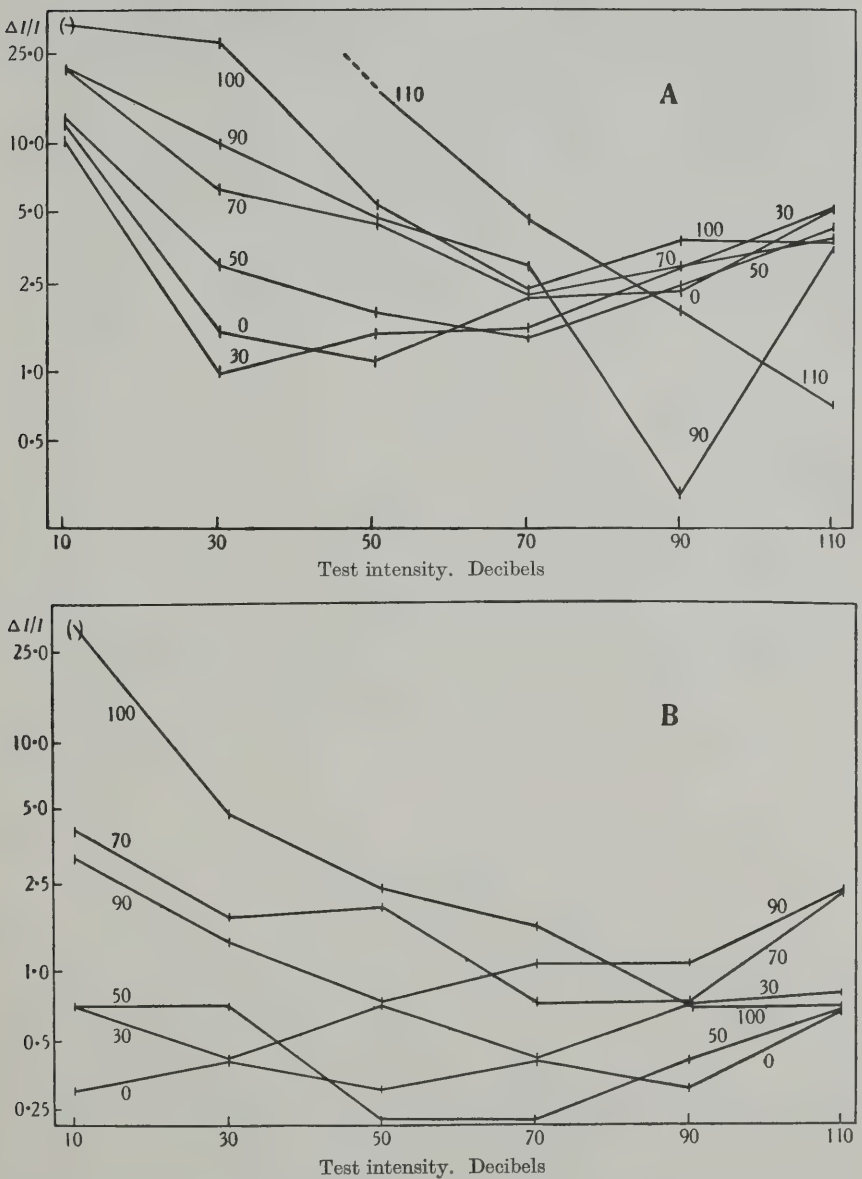


Fig. 5. Results of three observers A, B and C. The figures against each curve show the adaptation intensity in decibels above the absolute threshold. Frequency 1000 c.p.s.

instance, for example, if the 90 db. test had given a lower value of $\Delta I/I$ or the 70 db. test a higher value, the curve would have appeared typical. It is possible that such readings are in fact anomalous, due to wandering attention on the part of the observer, although he made no spontaneous acknowledgement at the time.

It may further be remarked that at low values of adaptation intensity (0, i.e. silence, and 30 db.), observers B and C (Fig. 5) show a very straight curve, while A gives a much higher value of $\Delta I/I$ at the lowest test intensities. It may be that observer A used a more certain criterion of the threshold than B or C at these low intensities.

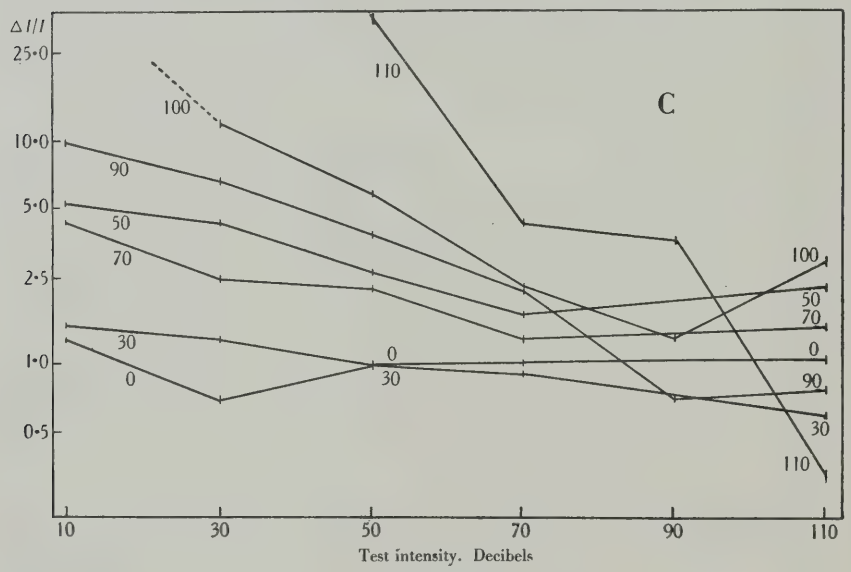


Fig. 5 C. See fig. 5 A, B for legend.

Such individual differences, however, are of minor importance by comparison with the general inference to be drawn from these curves; namely, that the region of the Weber's Law curve corresponding to the lowest values of $\Delta I/I$ shifts progressively towards those abscissae corresponding to high test intensities, when the ear is adapted to these high intensities. Indeed, the 110 db. adaptation curve reaches a very low value of $\Delta I/I$ at 110 db. test, and it may well be that it would rise again were it possible to test at higher intensities without causing serious pain to the observer.

It will be noted that with the higher intensities of adaptation, the

differential threshold becomes very large at low test intensities, and was on occasion greater than the deepest modulation which could be obtained with the apparatus. This is almost certainly due to the fact that at these intensities of adaptation, the absolute threshold rises, often as much as 30 db. (Rawdon-Smith, 1934, 1936). Thus in fact the observer may not hear the minima of the modulation at all. This speculation is borne out, moreover, by subsequent experiment, in which just this effect was demonstrated. It follows that, at 100 and 110 db. adaptation, the differential threshold reading obtained at test intensities below 15–20 and 30–40 db. respectively is no more than a measure of the absolute threshold. In consequence, such readings have been bracketed in the graphs A, B and C (Fig. 5).

All the above experiments were done at a fixed frequency of 1000 c.p.s. A further series was undertaken, in which the ear was adapted to a 1000 c.p.s. tone at 100 db., and tested at an intensity 30 db. above the respective thresholds at 100, 250, 500, 1000, 2000 and 4000 c.p.s. The results of this series can be plotted in graph form (Fig. 6 A, B and C) and show the value of $\Delta I/I$ obtained at the various frequencies, first with adaptation to silence, and then with adaptation to a 1000 c.p.s. tone at 100 db. In each case, the 1000 c.p.s. adaptation tone had no marked effect at lower test frequencies; it showed the expected effect at the same frequency, a comparable effect at 2000 c.p.s., and a slightly smaller effect at 4000 c.p.s. It is evident, therefore, that the effect of previous adaptation on the differential threshold is not confined to the adaptation tone frequency. This was also shown (Rawdon-Smith, 1934) to be true of the absolute threshold.

V. Conclusions

In attempting to draw conclusions from the results which we have summarized above, it appears relevant to effect some comparison between these data and those derived from related investigations in the auditory and the visual senses. To certain earlier work on the effect of intense stimulation on the absolute sensitivity of the human ear we have already referred (p. 124), and we may here note simply that, like the effects there discussed, the loss of differential sensitivity described in the present paper presents, primarily, the characteristic of lability. Indeed, this and other characteristics of the phenomenon lead us tentatively to ascribe it, as in the earlier case, in larger part to centrally mediated than to more peripheral changes; though as in the former case a peripheral aspect may also be present. If this is true, it would appear that

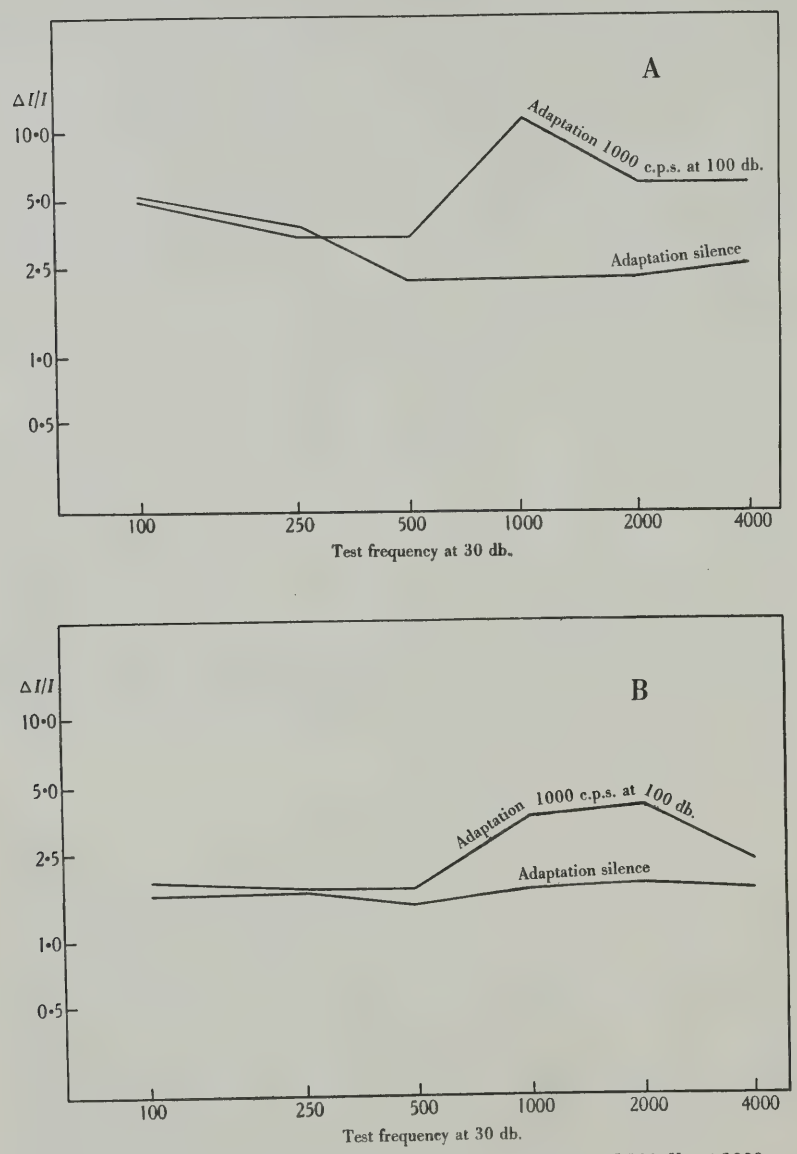


Fig. 6. Results of three observers A, B and C. Adaptation silence and 100 db. at 1000 c.p.s.; test 30 db. above threshold at the frequencies stated.

the differential threshold, like the absolute, is sensitive to what in an earlier paper (Rawdon-Smith, 1936) we have termed 'auditory inhibition'. Of this general phenomenon indeed, the loss of differential sensitivity described here may well be simply another aspect.

For the visual sense, Craik (1938) has recently shown an effect closely related to that which we have described; namely, a loss of differential sensitivity at low absolute intensities, after adaptation of the eye to a bright visual field. As in our case, differential sensitivity appears to be optimal when tested at an intensity close to that to which the sense organ has previously been subjected.

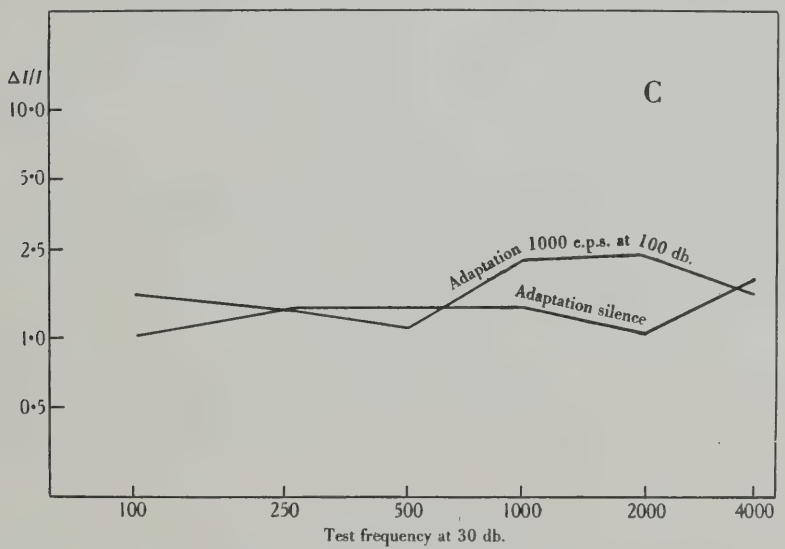


Fig. 6 C. See fig. 6 A, B for legend.

In many respects, indeed, the two senses—though in their mode of function so clearly different—here exhibit striking parallels. The significance of these for visual and auditory psycho-physiology, however, will be estimated with certainty only on the basis of further experimental work.

VI. Summary

A new technique for the assessment of the least perceptible increase in sound intensity has been evolved, using a method whereby a slow rhythmic change in voltage can be generated, and this made to modulate the output from a simple pure tone source. It is possible thus to estimate the differential threshold over a wide range of sound intensities, and

at any frequency which can be generated. It is further possible to measure the differential sensitivity at an intensity different from that to which the ear has been subjected for a previous period.

Using this technique, it has been found that the values of differential threshold differ greatly according to the state of adaptation of the ear: differential sensitivity is in general greatest when tested at the same intensity as that to which the ear has previously been adapted; it is less when the adapting intensity is greater or less than the test intensity.

It has further been found that adaptation at frequencies other than the test frequency does not have as great an effect as is observed when the adapting and test frequencies are the same, and often, indeed, no effect can be detected. Exceptions are found when the frequency of the adapting tone is equal to twice or four times that of the test tone.

Certain characteristics of these phenomena appear to indicate that the effect should be ascribed more readily to the intervention of central factors, than to modifications of sensitivity of the peripheral mechanism; this has also been shown to hold true of previously investigated effects on the ear's absolute sensitivity after stimulation. We believe, therefore, that like the latter, the effects discussed in this paper represent another aspect of the general phenomenon of auditory inhibition.

In part, the expenses of this investigation were defrayed by a grant from the Government Grant Committee of the Royal Society, to whom the authors wish to record their gratitude. They are also deeply indebted to Mr K. J. W. Craik, for assistance with the apparatus and as an observer.

REFERENCES

Craik, K. J. W. (1938). *J. Physiol.* xcii, 406.
Ewing, A. W. G. & Littler, T. S. (1935). *Brit. J. Psychol.* xxv, 284.
Fletcher, H. (1929). *Speech and Hearing.* New York: van Nostrand.
Knudsen, V. O. (1923). *Phys. Rev.* xxi, 84.
Rawdon-Smith, A. F. (1934). *Brit. J. Psychol.* xxv, 77.
—— (1935). *Nature, Lond.*, cxxxvi, 32.
—— (1936). *Brit. J. Psychol.* xxvi, 233.
—— (1938). *Theories of Sensation.* Camb. Univ. Press.
Rawdon-Smith, A. F. & Grindley, G. C. (1935). *Brit. J. Psychol.* xxvi, 191.
Riesz, R. R. (1928). *Phys. Rev.* xxxi, 867.

(*Manuscript received* 1 *April* 1939)

OBSERVATIONS RELATING TO THE THRESHOLD OF A SMALL FIGURE WITHIN THE CONTOUR OF A CLOSED-LINE FIGURE

By K. J. W. CRAIK and O. L. ZANGWILL

From the Cambridge Psychological Laboratory

I. *Introduction* (pp. 139–140).
II. *Preliminary experiments* (pp. 140–143).
(1) *Apparatus and method* (pp. 140–142).
(2) *Procedure* (p. 142).
(3) *Subjects* (p. 142).
(4) *Results and discussion* (pp. 142–143).
III. *Final experiments* (pp. 143–148).
(1) *Object and method* (pp. 143–144).
(2) *Apparatus* (pp. 144–145).
(3) *Procedure* (p. 145).
(4) *Results* (pp. 145–147).
(5) *Discussion* (pp. 147–148).
IV. *Summary and conclusions* (p. 149).
References (pp. 149–150).

I. Introduction

Koffka, in his discussion of the problem of shape ((12), pp. 129–60), calls attention to the fact that if a line forms a closed or almost closed figure it will tend to be perceived not as a mere line upon a homogeneous background but as a surface figure bounded by a contour ((12), p. 150). A figure of this kind he therefore calls a *closed contour figure*. Koffka then goes on to predict that the threshold of a variable is likely to be higher when projected upon the surface of a closed contour figure than when projected outside the bounding contour. This prediction is based partly on general grounds and partly on the results of earlier experiments (7, 8) relating to differences of threshold on figure and ground.

Koffka points out that his prediction might readily be put to the experimental test. He suggests to this end a modification of a method originally devised by Hempstead (11). The proposed method consists essentially in projecting a small figure upon the contoured surface and

arranging an episcotister between the latter and the projection lantern. The differential threshold of the small figure could then be determined by suitable adjustment of the variable episcotister aperture. "If then", writes Koffka, "the little figure required a greater episcotister opening in order to become visible inside than outside the contour, we should have proved a greater cohesiveness of the enclosed area as compared with its surroundings, which would make it more difficult to produce a new figure on it" ((12), p. 150). "Unfortunately", he concludes, "this experiment has never been made, although from two similar experiments, one by Gelb and Granit and the other by Granit, our assumed result seems predictable."

The purpose of the present study is to remedy the deficiency emphasized by Koffka and thereby to submit his theoretical prediction to the experimental test.

II. Preliminary experiments

(1) *Apparatus and method*

A preliminary test utilizing the Hempstead technique (described above) was carried out which gave strong indications of a positive result. This technique is however unsatisfactory for a number of reasons: In the first place, a great deal of time is consumed in obtaining an adequate number of threshold determinations from each subject. In the second place, the noise engendered in varying the episcotister aperture during rotation has a somewhat distracting effect on the subject. In the third place, it is difficult to vary the aperture during rotation in a regular fashion. Consequently, we agreed to abandon the Hempstead technique and to determine the threshold of the variable figure with the aid of an iris diaphragm. The apparatus employed in these preliminary experiments was originally devised by Grindley (10) for group demonstration of Weber's Law, but it could readily be adapted for our present purpose.

The apparatus (Fig. 1) consisted of a lamp (L) placed at a distance of 1·8 m. from a white cardboard screen (SS) which it illuminates directly through the window (W). Some light from L is reflected on to the screen by the mirror (M) and this light passes through the lens (O). A small rectangle, 2 × 3 mm., was drawn in black ink upon an ordinary glass notice plate (P) interposed between L and M. An image of the rectangle is thrown on the screen, and the difference between its brightness and that of the background can be varied by means of the iris diaphragm (I) on the lens. The diaphragm was calibrated directly for $\delta I/I$ on a scale attached to the exterior of the instrument. The brightness of L was adjusted to a constant value for the purposes of the experiment. The value chosen for direct illumination of the screen was 0·50 foot-

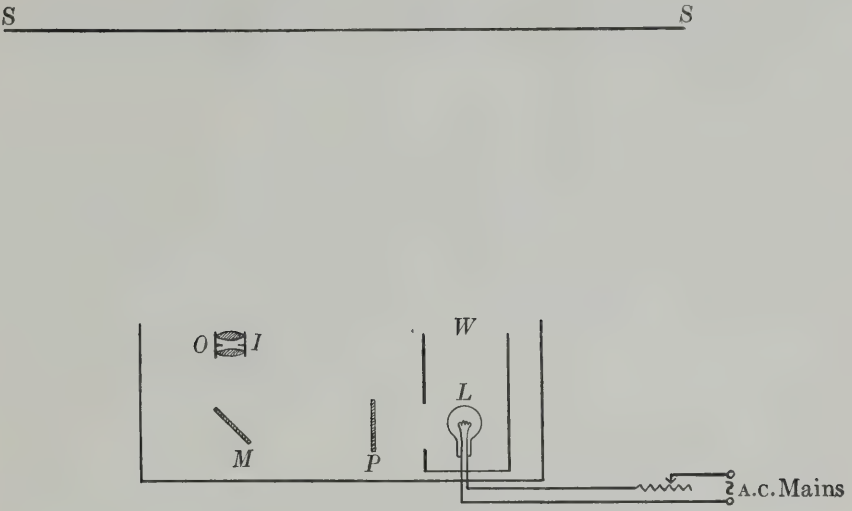


Fig. 1.

candles.[1] It is obvious that enlargement of the diaphragm aperture under these conditions will lead to a slight increase in total illumination of the screen. This is not, however, perceptible as a change in background illumination but as a difference between the brightness of the figure and the ground. The changes made by this method give the threshold as $\frac{\delta I}{I+\delta I}$ which is virtually equivalent to $\delta I/I$ for the small ratios of $\delta I : I$ employed.

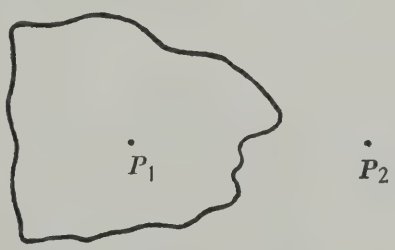


Fig. 2. Original size reduced by 0·5.

The screen SS was a white cardboard sheet, 25×20 in., suspended in such a way that it could be shifted horizontally in either direction. An irregular figure (Fig. 2) was drawn in black crayon at about the centre of the screen. For purposes of fixation, a small black point (P_1) was marked within the contour of the figure such that when the image of the rectangle is projected inside the border, this point approximately marks

[1] The illumination conditions were of course maintained constant for all positions of the variable figure. Subsequent determinations in which a higher value for the screen was adopted gave results identical with those reported below.

what would be the point of intersection of its diagonals. A second point (P_2), outside the border of the larger figure, marks the fixation point under conditions in which the image of the rectangle is arranged to fall outside the contour.

(2) *Procedure*

The subject was seated to the right of the apparatus in such a manner that his head and the lamp L were approximately level. The threshold was then determined in the following way: With the image of the rectangle projected on the contour figure, five readings of the threshold for 'disappearance' followed by five for 'appearance' of the variable figure were obtained. The screen was then shifted so as to bring the image of the rectangle on to a specified area of the screen outside the border of the contour figure. A further set of five readings for 'disappearance' and five for 'appearance' was then obtained in an exactly similar way. The screen was then shifted back to its original position and ten further readings obtained as before. Finally, a set of ten determinations was secured with the image of the rectangle falling once again in its previous position outside the contour.

Special precautions were exercised with reference to (*a*) the notorious lack of definition of the differential threshold; (*b*) the rate of appearance and disappearance of the variable; and (*c*) the possible effects of practice and fatigue. Thus in the first place, we instructed each subject to adopt and maintain as distinctive a criterion of the threshold as he was able. In the second place, we endeavoured to operate the diaphragm at a constant rate during any one given determination, but to vary the rate slightly as between the successive determinations. We also instructed the subjects to go as little as possible by time interval in effecting their judgements as this would not remain constant for all trials. In the third place, we gave rest pauses wherever necessary in order to reduce to a minimum interference by after-imagery. And in the fourth place, as already described, we distributed the trials under the two sets of conditions to offset any differential effect of practice upon discrimination.

(3) *Subjects*

Twenty subjects, the majority of whom were either research workers or students in psychology, took part in the experiment. With two exceptions, all were unaware of the purpose of the investigation.

(4) *Results and discussion*

The mean value of the differential threshold of the small figure when projected within the contour of the larger figure was found to be higher than the mean threshold of the same figure when projected outside the contour in the case of nineteen out of the twenty subjects. The mean value of the difference between the two sets of readings for all subjects was found to be statistically significant. This result provides a clear confirmation of Koffka's predicted finding. It does not, however, necessarily establish the validity of Koffka's theoretical position. Our result

might equally well be due to the operation of some differential contrast effect. Further experiment is therefore required to establish the origin of the phenomenon under discussion.

III. Final experiments

(1) *Object and method*

A final, and we hope decisive, series of experiments was planned to determine whether our preliminary result was due to the intrinsic cohesion of a contour figure or whether it derives from some process allied to contrast.

Our method is described below:

Two series of drawings were prepared. Those forming the *Figural* (*F*-) *Series* are shown in Fig. 3; those forming the *Control* (*C*-) *Series* in Fig. 4. The procedure involved a series of threshold determinations of a small patch of variable intensity at the centre of the square shown in Fig. 3*d* or at a strictly corresponding position in the case of each of the remaining figures shown below.

Fig. 3.

Fig. 4.

It is to be expected on Koffka's view that the threshold of the variable would show *either* a sudden rise on complete closure of the square (*F*-series, stage *d*) *or* a progressive rise with increasing figural closure (*F*-series, stages *c*, *d*). If the first alternative obtains, one would expect no corresponding terminal rise of threshold in the case of the *C*-series; if the second obtains, one would predict consistently higher threshold values for stages *c* and *d* in the *F*-series as compared with stages *x*, *y*, and *z* in the *C*-series.

If, on the other hand, the entire effect is due to contrast, no such differences as the above would be expected, since the size and distance of the lines in corresponding members of the two series are identical.

In conclusion, to check the analogy between the present conditions and those obtaining in the preliminary experiment, a limited set of values of the threshold inside and outside the contour of the square was secured.

(2) *Apparatus*

The apparatus (Fig. 5) consisted of a set of nine identical white card screens (S_1), each 20×13 in., made up in book form, hinged at one side and fitted with a pin to register them accurately at the other side. In the centre of each screen was a very

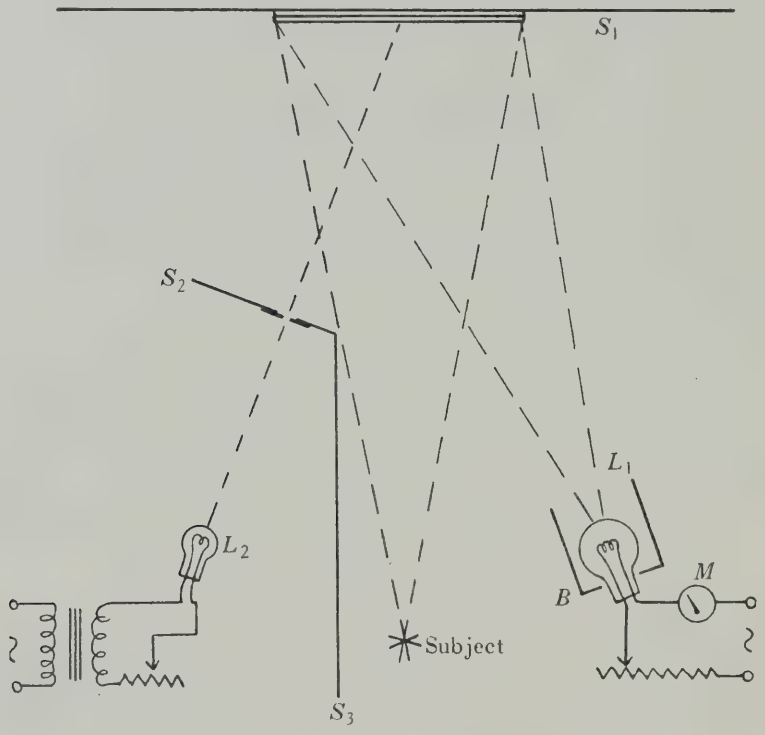


Fig. 5.

small fixation point. The first screen was otherwise blank; the others had one or more Indian ink lines each 1 in. long and $\frac{1}{2}$ in. from the fixation point, as shown in Figs. 3 and 4 above. The last member of the *C*-series consisted of two lines having twice the width of those in the other fields and consequently equivalent in black area to a four-line figure.

The screen was illuminated by a 12 V. 24 W. bulb (L_1), supplied from the A.C. mains through a resistance. Its illumination was sufficiently uniform over the field and was arranged to be 1·5 foot-candles. The current through the lamp was measured continuously by an ammeter (M) and found constant to 2%. The lamp was shielded from the subject by a box (B). On the left of the subject, who sat in a fixed chair with

his eyes at a distance of 50 in. from the screen, was a 4 V. bulb (L_2) with a compact filament casting on the screen S_1 a sharp bright shadow of a small rectangular aperture in the screen S_2. The position of the bright shadow could be adjusted to coincide exactly on each screen by having the aperture in a movable metal slide on S_2. This bulb was supplied by a transformer from the mains and its brightness was varied by means of a 25 ohm series resistance having a very smooth and silent motion; this gave a range of illumination on the screen from 0·002 to 0·15 foot-candles. The resistance and the bulb were screened from the subject by S_3. The brightness of the variable illumination on the screen was measured by a Holophane Lumeter and checked by an Everett Edgcumbe Luxometer, and the resistance calibrated directly in foot-candles. Although varying the brightness by a series resistance alters the colour of the light considerably, its total brightness is never more than 10% and its threshold brightness 2% or less, so that the change of colour is lost in the main illumination.

The experiments were carried out in a dark room with a very small lamp continuously illuminating the resistance scale and the experimenter's record of the readings.

The rectangular screen subtended angles of $20° \times 15°$ at the subject's eye; the complete figure, 1° 12′ sq. and the test patch, $24' \times 10'$.

(3) *Procedure*

Eight subjects were employed in the main experiment. These comprised five research workers and three students in psychology.

In the case of five of the subjects, the determinations were made in the following order: Blank field; *F*-series in the sequence *a*, *b*, *c*, *d*; blank field; *C*-series in the sequence *w*, *x*, *y*, *z*; blank field. In the case of the remaining three subjects the whole order of presentation was inverted.

Ten threshold values were determined for every figure, five 'appearing' and five 'disappearing'. A total of ten readings was likewise taken in the case of every determination on the blank field.

The subject was instructed to say 'stop' when the brightness of the variable reached threshold value. In order to obviate judgements based on time interval, the experimenter in changing the brightness endeavoured to vary both the speed and the starting-point of each determination.

(4) *Results*

The mean threshold value of each subject in the case of the blank field and of each of the two series of test fields is given in Table I. The symbols (*s*) and (*i*) stand for 'straight' and 'inverted' order respectively.

The ratio of the threshold on each field to the mean threshold taken as unity is shown in Table II.

It will be seen from Table II that the mean differential threshold values at equivalent stages of the two series are identical. This result is also given graphically by Fig. 6.

The mean variation, indicative of degree of scatter, is 0·17.

Table I

Subject	Field: Blank	*a*	*w*	*b*	*x*	*c*	*y, z*	*d*
R.H.T. (*s*)	0·044	0·061	0·049	0·066	0·065	0·072	0·075	0·079
R.C.O. (*s*)	0·019	0·032	0·028	0·041	0·036	0·037	0·035	0·035
M.H.H. (*i*)	0·041	0·041	0·067	0·060	0·061	0·058	0·062	0·065
M.G.C. (*s*)	0·081	0·078	0·105	0·094	0·130	0·100	0·122	0·104
A.T.W. (*s*)	0·066	0·064	0·088	0·079	0·095	0·093	0·100	0·093
O.L.Z. (*s*)	0·023	0·029	0·041	0·027	0·039	0·035	0·040	0·034
K.J.W.C. (*i*)	0·029	0·035	0·036	0·044	0·042	0·047	0·046	0·052
J.G.F. (*i*)	0·081	0·086	0·092	0·096	0·100	0·101	0·108	0·108

Table II

Subject	Field: *a*	*w*	*b*	*x*	*c*	*y, z*	*d*
R.H.T.	1·4	1·1	1·5	1·5	1·6	1·7	1·8
R.C.O.	1·7	1·5	2·1	1·9	1·9	1·8	1·8
M.H.H.	1·0	1·6	1·5	1·5	1·4	1·5	1·6
M.G.C.	0·9	1·2	1·2	1·6	1·2	1·4	1·3
A.T.W.	1·0	1·3	1·2	1·4	1·4	1·5	1·4
O.L.Z.	1·3	1·8	1·2	1·7	1·5	1·7	1·5
K.J.W.C.	1·2	1·2	1·5	1·5	1·6	1·6	1·8
J.G.F.	1·1	1·1	1·2	1·2	1·2	1·3	1·3
Mean	1·2	1·4	1·4	1·5	1·5	·6	1·6

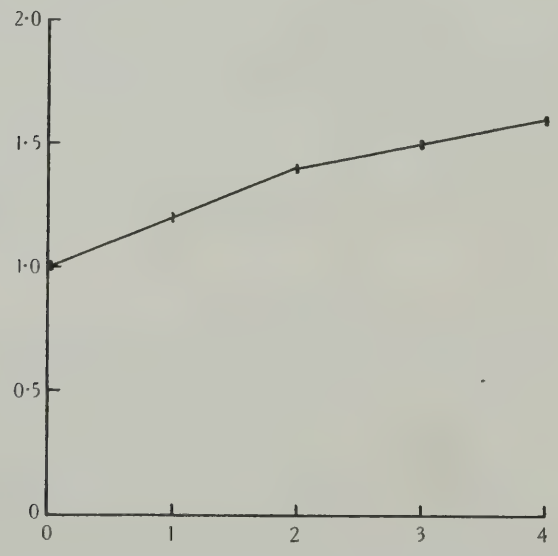


Fig. 6. Ordinate: mean ratio of threshold near lines to threshold on blank field where latter is taken as unity. Abscissa: number of lines in the field.

An additional set of threshold determinations inside and outside the contour of the square is presented in Table III. Each value in the first column of this table constitutes a mean threshold value of the variable in the position shown in Fig. 3 *d*. Each value in the second column constitutes the mean for a set of threshold determinations with the variable $\frac{1}{2}$ in. to the left of one vertical side of the square and a set with the variable $\frac{1}{2}$ in. to the right of the other vertical side. The ratio 'inside'/'outside' is shown in the third column.

Table III

Subject	Inside (*i*)	Outside (*o*)	*i*/*o*
K.J.W.C.	0·041	0·028	1·5
O.L.Z.	0·041	0·029	1·4
M.D.V.	0·051	0·036	1·4

(5) *Discussion*

The results of our final series of experiments clearly show that neither expectation set up by Koffka's cohesion hypothesis is satisfied. In the first place, there is a *flattening* at the final stages of the curve for the *F*-series. In the second place, the curves for the two series actually coincide throughout. We thus seem bound to conclude that the rise in threshold noted under our experimental conditions cannot be due to the intrinsic cohesiveness of an enclosed area.

Our conclusion finds a measure of support in the results of some recent experimental studies. Thus Fry & Robertson, in their re-examination of the alleged influence of figural properties upon hue and brightness, are led to conclude that "...the observed effects instead of being due to the mode of perception are due rather to neglected factors in the conditions of ocular stimulation" ((6), p. 424). Special attention was devoted by these authors to the problem of the differential brightness threshold on figure and ground. "We have avoided the difficulties inherent in Gelb and Granit's method", they claim, "by using a pattern ...whose parts can be seen interchangeably as figure and ground, without involving any change in the stimulus conditions, and in such a case we have found that there is no significant difference between the threshold on figure and on ground" ((6), p. 434). This result, when taken in conjunction with our own findings, would seem to indicate that the indubitable rise in threshold observed in our own experiments, as also in certain of those of Fry & Robertson ((6), pp. 426–7) and Fry & Bartley ((5),

pp. 415–16), must be attributed not to "differences in perceptual relationships" but to "changes produced in the peripheral physiological processes" ((6), p. 426).

We must concede that very little is known at present as to the precise nature of these peripheral changes. But there is already ample evidence of very marked effects on the differential brightness threshold due to a dark surrounding field (Abribat (1), Cobb (2), Lythgoe (13), Lythgoe & Tansley (14)). These findings, together with those of Craik (3, 4), indicate that a somewhat dark adapted retina is ill-equipped to discriminate differences at a fairly high illumination. It is tempting, and not altogether speculative, to suppose that a dark figure (or 'surround') causes some degree of dark adaptation in neighbouring areas of the retina akin to that produced by subjection of the *whole* retina to a period of comparative darkness. For as one of us (Craik) has pointed out elsewhere, "...spatial and temporal proximity of an adapting field appear in general to have very similar effects" ((4), p. 252). We may note also that Schouten (16) and Schouten & Ornstein (17) come to a similar conclusion.[1]

The general implication of all this work would seem to us to be that the threshold effects discussed above cannot profitably be treated as functions of 'organization' in Koffka's sense of the term. In the first place, as we have endeavoured to indicate, other and more appropriate lines of explanation are available. In the second place, there seem to us intrinsic reasons why the threshold technique is fruitless when applied to the investigation of figural properties. It is our belief that the processes underlying differential brightness sensitivity are peripheral in origin. They must in consequence be treated by the organism at the lowest levels of functional integration. Form perception, and the apprehension of figural properties, we believe to be founded on activities located at higher integrative levels (cf. Parsons ((15), pp. 45–60)). Nevertheless, there seems to us no good reason why the high level processes which find their expression in the emergence of form need necessarily obscure those whose determinants are purely peripheral. Thus in cases such as that reported in the present study, it would seem to us idle to attempt to establish the phenomenal reality of an attribute of shape by means of the threshold technique. Prof. Koffka's theory of perception, despite its brilliant synthetic qualities, seems to us severely limited by its author's persistent refusal to countenance any principle of 'levels' in the perceptual process.

[1] Granit (9) has recently been able to show that the differential brightness sensitivity of the retina, as measured by its electrical response, is decreased (at moderate illuminations) by previous dark adaptation.

IV. Summary and conclusions

1. The threshold of a small rectangle of variable brightness is significantly higher when projected inside the contour of a larger irregular closed figure than when projected outside the contour. A similar result is obtained in the case of a small patch of variable intensity projected inside and outside the contour of a square.

2. These results substantiate Koffka's prediction that the differential threshold of a variable will be raised by its projection on a contoured surface. They do not, however, provide unequivocal support for his theory of the intrinsic cohesiveness of a closed contour figure.

3. The outcome of a crucial experiment shows that there is a progressive rise in threshold of a variable as black lines of specified dimensions are added to the field in which the brightness change occurs. It is however pointed out that the rise is *identical* whether the lines are arranged to build a closed square or merely to constitute two sets of parallel lines.

4. It is concluded from these results that the observed rise in threshold cannot depend upon Koffka's factor of intrinsic cohesion. An alternative interpretation, more in harmony with the known facts of visual adaptation, is tentatively advanced. Certain broader implications of the results are briefly discussed.

The writers are profoundly indebted to Prof. K. Koffka for the interest which he has taken in these studies. They alone, however, are responsible for the theoretical interpretations presented above, to which Prof. Koffka does not himself subscribe. Gratitude is also due to Prof. F. C. Bartlett for his encouragement and help throughout.

REFERENCES

(1) Abribat, M. (1935). *Compte Rendu de la Réunion de l'Institut d'Optique*, April.
(2) Cobb, P. W. (1916). *J. exp. Psychol.* I, 540.
(3) Craik, K. J. W. (1938). *J. Physiol.* XCII, 406.
(4) —— (1939). *Brit. J. Psychol.* XXIX, 252.
(5) Fry, G. A. & Bartley, S. H. (1935). *Amer. J. Physiol.* CXII, 414.
(6) Fry, G. M. & Robertson, V. M. (1935). *Amer. J. Psychol.* XLVII, 424.
(7) Gelb, A. & Granit, R. (1923). *Z. Psychol.* XCIII, 83.
(8) Granit, R. (1924). *Skand. Arch. Physiol.* XCV, 43.
(9) —— (1937). *Nature, Lond.*, CXXXIX, 719.

(10) Grindley, G. C. (1930). *J. gen. Psychol.* iv, 427.
(11) Hempstead, L. (1900). *Amer. J. Psychol.* xii, 185.
(12) Koffka, K. (1935). *Principles of Gestalt Psychology.* London: Kegan Paul & Co.
(13) Lythgoe, R. J. (1932). *Med. Res. Coun. Spec. Rep. Ser.*, no. 173.
(14) Lythgoe, R. J. & Tansley, K. (1929). *Proc. roy. Soc.* B, cv, 60.
(15) Parsons, Sir John (1927). *An Introduction to the Theory of Perception.* Camb. Univ. Press.
(16) Schouten, J. F. (1939). *International Commission on Illumination,* iv, Paper 3.
(17) Schouten, J. F. & Ornstein, L. S. (1939). *J. opt. Soc. Amer.* xxix, 168.

(*Manuscript first received* 1 *April* 1937)

SOME FACTORS OF PERSONALITY.

A FURTHER ANALYSIS OF SOME OF WEBB'S DATA

By H. A. REYBURN and J. G. TAYLOR

From the Psychological Laboratory of the University of Capetown

I. *Introduction* (pp. 151–154).
II. *Further analysis of Webb's data* (pp. 154–165).
(1) *Traits selected* (pp. 154–155).
(2) *Method employed* (pp. 155–157).
(3) *Results obtained* (pp. 157–161).
(4) *Conclusions as to qualities isolated* (pp. 162–165).
References (p. 165).

I. Introduction

One of the main problems in the analysis of temperament and character is that of finding a suitable framework in which to set the results of observation and experiment. A great deal of activity has been shown by a variety of authors in attacking this problem, but unfortunately no general agreement has been reached. A factor proposed by one author may be accepted by another, with or without modification. Thus, for example, the factor '*w*' which Webb isolated in his classic study on "Character and Intelligence" (11), has been recognized, with some alterations, by Garnett (5) and by Cattell (3). The factor '*c*', isolated also from Webb's material, was named 'cleverness' by Garnett, and stands obviously in close relation both to the Surgence-Desurgence factor of Cattell and to the Dominance-Submission factor of the Bernreuter Personality Inventory (1), although Spearman (8) would prefer to regard it as the inverse of perseveration. But on the other hand, although single factors are common to several schemes, they are supplemented in different ways by different authors, and there is no common framework which everyone can accept.

This difficulty arises in part from a weakness that characterizes more than one of the methods of analysis that have been used. Procedures based on the technique devised by Spearman involve the selection of a relatively small group of traits such that their intercorrelations will satisfy the tetrad equation criterion. A single factor is then postulated, and an attempt made to identify it. Then, the influence of this factor

being partialled out, a new grouping is made, and a further factor is extracted. A result of this method is that when a factor is discovered its nature is assumed to be fixed, and is not reconsidered in the light of the factors that are isolated afterwards. That is to say, a number of reference axes are laid down, one after another, each being fixed by itself, no attempt being made to discover the best position for the co-ordinate frame as a whole. And yet it is obvious that the initial positions of these axes are only provisional, and may have to be altered in the light of further data.

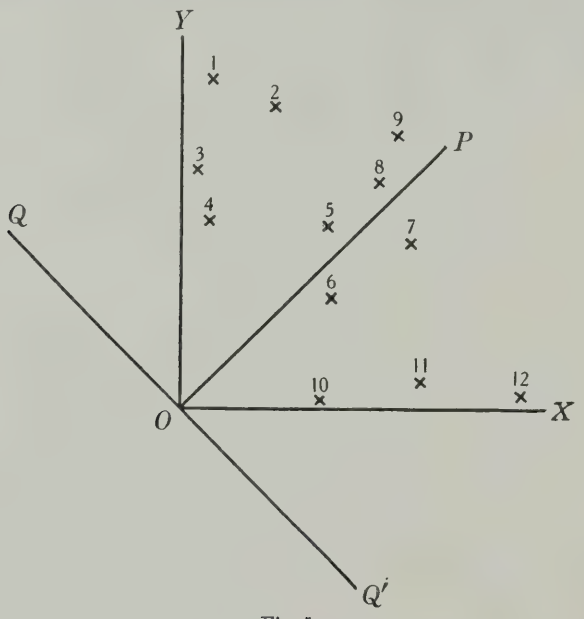


Fig. 1.

For example, suppose there are twelve traits having two independent common factors such that the factor loadings are either positive or zero, and let these two factors be represented by the co-ordinate axes OX and OY in Fig. 1. Suppose further that the technique we have been discussing is applied to this material, and traits 5, 6, 7, 8 and 9 are considered as a group for preliminary analysis. Within the limits of the probable error, the correlations between these traits satisfy the tetrad criterion, and consequently a single common factor is announced. It is indicated in the diagram by the axis OP, drawn from the origin through the points to fit them most closely. But when the remaining traits are considered, it is clear that this axis does not represent a single primary quality. When a second axis is fitted to explain the remaining intercorrelations of the

whole group of traits, it must be drawn at right angles to *OP* in order to preserve the independence of the factors. Let this axis be represented by the line *OQ*. It follows at once that the loadings of about half the traits are negative, and that is in contradiction with our postulate that all factor loadings are either positive or zero. The difficulty is entirely removed if the axes are rotated from *OP*, *OQ* to the positions *OX*, *OY*. And when this is done, the supposed factor *OP* disappears.

This example is given to illustrate a general principle. The particular assumption made for convenience of exposition, namely, that there are no negative factor loadings, is not intended to apply beyond this imaginary case, and is merely used to represent *any* relevant psychological considerations. The underlying principle is that the final positions of all the axes of reference can be determined only when they are all considered together. If the significance provisionally given to the factors first discovered renders it difficult, or even impossible, to give a reasonable psychological interpretation to others that are later found necessary to complete the explanation of the material, then the axes as a whole must be rotated in such a way that a meaning can be given to them all.

Even when procedures more powerful than those derived from the tetrad equation criterion are employed, this consideration is not always given due weight. The point is well expressed, not for the first time, by Thurstone (10).

> One of the most common characteristics of factor studies is the computation of a factorial matrix and the attempt to make a psychological interpretation without rotating the reference frame. It does not matter whether the factorial matrix has been produced by the centroid method or by the principal axes or components. In either case the configuration must be rotated before the factors can be even expected to have psychological meaning. One cannot expect to find several meaningful factors by any of the current methods of factoring a correlational matrix without rotation of the reference frame. (p. 90.)

Perhaps an illustration will make this clearer. Taking from Brigham a table of the intercorrelations of marks obtained by officers and cadets of the U.S. Military Academy in eleven subjects, viz. English, History, Economics, Law, Engineering, Ordnance, Physics, Chemistry, Mathematics, Surveying and Drawing, Flanagan (4) analyses it by Hotelling's method. Only the first two factors need concern us here. Taking them as they stand, the author is driven to describe them thus:

> We might call the first factor a general-achievement factor. None of our more elemental concepts seems to fit it particularly well. Evidence from other studies of achievement would lead one to believe that a major part of this may be considered to

be due to industry, which is itself a complex of such things as interest, purpose, temperament, perseverance, motivation, etc. Probably Spearman's '*g*' may also be thought of as contributing quite substantially to this first factor.

The second factor... has its largest factor-loading in drawing, with surveying and mathematics next in order. The largest negative factor-loading is for history, followed by English, economics and law. This second factor seems, then, to represent a difference between spatial ability, perhaps combined with some numerical ability and a certain amount of manual dexterity, and verbal ability, with possibly some memory involved. As a further speculation we suggest that it might be linked with some physiological or temperament factor which gave mental activities a singleness of direction and resistance to change very similar to the inertia of a moving body. (p. 33.)

Could the case for rotation be stated better? If the loadings of these two factors are plotted, and the axes then turned through a negative angle of 29° or 30°, the picture changes entirely. We now get one factor with negligible loads on history, economics, English and law, and with significant loads on engineering, ordnance, physics, chemistry, mathematics, surveying and drawing—obviously not an impossible situation to interpret; and a second factor with a negligible load on drawing, with heavy loadings on English, history, economics and law, and with intermediate loadings on the other subjects—again a factor of reasonably plain meaning. Further rotation in a larger number of dimensions might improve the situation still more, but enough has been said to show the undesirability of trying to give a psychological interpretation to the products of factorial analysis as they come straight from the mathematical mill.

II. Further analysis of Webb's data

(1) *Traits selected*

In view of these considerations it struck us that before making use of material gathered by ourselves it might be advisable to re-examine the data collected by Webb. Webb's material, however, is very extensive, and the factorial analysis of the whole of his correlation matrix seemed to us to be too large a task until we had more guiding posts than are at present available. Accordingly we selected nineteen of Webb's traits for further treatment. In his analysis '*g*' plays a fairly large part, but we had no desire to examine this factor in the present enquiry. We therefore began by selecting the eight traits by means of which Webb attempted to establish the existence of '*w*', since from his figures it is clear that '*g*' plays a very small part in producing their intercorrelations. To these we added eleven other traits which seemed to have a bearing on emotional disposition and behaviour. And as the results obtained by Webb for his

groups of men students are more extensive and probably more reliable than those for his groups of schoolboys, we have employed the former throughout.

The following list gives the names of the traits in question, and the numbers attached to them by Webb in his schedule:

1. General tendency to be cheerful (as opposed to being depressed and low-spirited).
2. Tendency to *quick* oscillation between cheerfulness and depression (as opposed to permanence of mood).
3. Occasional liability to extreme depression.
4. Readiness to become angry.
5. Readiness to recover from anger.
6. Occasional liability to extreme anger.
8. Degree of sense of humour.
11. Eagerness for admiration.
15. Fondness for large social gatherings.
17. Impulsive kindness (to be distinguished from no. 18).
18. Tendency to do kindnesses on principle.
19. Degree of corporate spirit (in whatever body interest is taken, e.g. college, school, country, native place, etc.).
20. Trustworthiness (keeping his word or engagement, performing his believed duty).
21. Conscientiousness (keenness of interest in the goodness and wickedness of actions).
23. Readiness to accept the sentiments of his associates.
27. Degree of 'tact' in getting on with people.
31. Degree of bodily activity in pursuit of pleasure (games, etc.).
33. Tendency *not* to abandon tasks in the face of obstacles.
34. Tendency *not* to abandon tasks from mere changeability.

(2) *Method employed*

Turning now from the material, we come to the problem of method. Of the various devices available we have adopted that of Thurstone (9). Thurstone himself has stated clearly the limitations and weaknesses of his centroid method of factorizing a correlation matrix, but it is the simplest procedure, and we are not satisfied that the advantages claimed for such other methods as those of Kelley, Hotelling or Burt (2) are sufficient to compensate for their greater complexity. The positions of the axes as reached by the different methods are not identical, but if the different frames of reference are each orthogonal, it is possible to pass from one to another by suitable rotations. We have already seen that the axes have to be rotated in order to give them psychological significance. There is therefore no special virtue in any particular starting point. The difference

Table I. *Correlation matrix extracted from Webb's Table IV.*

Decimal points are omitted.

	1	2	3	4	5	6	8	11	15	17	18	19	20	21	23	27	31	33	34
1		−27	−53	−10	43	−16	68	12	51	20	19	42	07	−08	−01	21	36	23	06
2	−27		48	42	−27	38	01	31	10	−03	−43	−06	−43	−29	25	−12	05	−52	−48
3	−53	48		24	−33	21	−35	04	−30	−18	−26	−22	−14	−09	−06	−04	−27	−49	−24
4	−10	42	24		−63	70	11	65	49	−21	−81	−07	−55	−58	−45	−38	24	−21	−31
5	43	−27	−33	−63		−41	22	−32	02	41	48	32	45	33	26	48	−02	15	31
6	−16	38	21	70	−41		−13	53	26	−05	−34	03	−38	−39	−31	−32	23	−18	−30
8	68	01	−35	11	22	−13		12	55	34	26	46	01	−07	−19	42	30	23	−09
11	12	31	04	65	−32	53	12		36	03	−41	−12	−51	−51	−34	−26	25	−27	−31
15	51	10	−30	49	02	26	55	36		29	−04	51	−37	−49	−35	04	55	−13	−31
17	20	−03	−18	−21	41	−05	34	03	29		50	34	30	26	21	34	10	01	06
18	19	−43	−26	−81	48	−34	26	−41	−04	50		50	76	71	27	48	04	35	36
19	42	−06	−22	−07	32	03	46	−12	51	34	50		28	22	04	36	54	21	09
20	07	−43	−14	−55	45	−38	01	−51	−37	30	76	28		80	16	33	−13	39	53
21	−08	−29	−09	−58	33	−39	−07	−51	−49	26	71	22	80		24	27	−22	39	52
23	−01	25	−06	−45	26	−31	−19	−34	−35	21	27	04	16	24		12	−09	−17	00
27	21	−12	−04	−38	48	−32	42	−26	04	34	48	36	33	27	12		−02	10	13
31	36	05	−27	24	−02	23	30	25	55	10	04	54	−13	−22	−09	−02		−01	−12
33	23	−52	−49	−21	15	−18	23	−27	−13	01	35	21	39	39	−17	10	−01		63
34	06	−48	−24	−31	31	−30	−09	−31	−31	06	36	09	53	52	00	13	−12	63	

between the various systems is, however, too large a problem to discuss here, and merits separate consideration.

The correlation matrix which we have analysed is given in Table I.

One of the main problems in the application of any method of factorial analysis to fallible data is that of deciding how many factors should be extracted. If the material is free from error, the rank of the matrix can be accurately determined, and hence the number of factors to be extracted is known. But when the correlations are subject to error, the residuals may not vanish completely until as many factors have been taken out as there are variables in the correlation table. In other words, the common factor space cannot be separated from the error space, and the analysing mechanism will continue to operate until both have been exhausted.

It seems clear that the process should continue so long as the residuals are too large to be accounted for by chance, and that, on the other hand, in accordance with the principle of parsimony, it should stop when they have been so far reduced that the probable errors of the correlation coefficients are sufficient to account for them. In order to decide when this point had been reached, we thought it desirable to divide the residuals at each stage by the probable errors of the original coefficients, and to plot a frequency distribution of the results, disregarding signs. Treating this curve as one-half of a total distribution, we compared it with a normal curve covering the same area and having a standard deviation of 1·4825. If the fit is poor, it follows that something more than the operation of pure chance is required to account for the residuals.

(3) *Results obtained*

After three factors had been extracted, the residuals by this test proved to be too large, and after the fourth factor they were too small. The values of P yielded by Pearson's test for goodness of fit were, for the third factor residuals 0·17, and for the fourth factor residuals 0·07. As it is not possible to extract three and a half factors, we decided to retain the fourth, and to recognize that a certain amount of error space is mixed up with the common factor space.

The matrix M, derived by the centroid method from Table I, is presented in Table II. The first four columns give the factor loadings of the nineteen traits analysed, and the fifth column, $h2$, gives the communalities of the traits.

As it stands this matrix has no psychological significance, and its axes require to be rotated until they represent recognizable psychological factors. As a starting point we took the factor 'c' already isolated from

Table II. *Centroid matrix* M

	I	II	III	IV	h^2
1	0·344	0·665	−0·117	−0·320	0·667
2	−0·520	0·015	0·436	0·207	0·504
3	−0·415	−0·374	0·337	0·221	0·475
4	−0·838	0·309	−0·308	0·168	0·921
5	0·672	0·155	0·196	−0·105	0·525
6	−0·618	0·261	−0·178	0·375	0·622
8	0·231	0·755	0·063	−0·210	0·671
11	−0·584	0·403	−0·171	0·025	0·533
15	−0·188	0·875	0·061	0·014	0·805
17	0·373	0·373	0·261	0·210	0·390
18	0·837	0·077	0·133	0·287	0·807
19	0·336	0·608	0·074	0·414	0·659
20	0·788	−0·186	−0·121	0·340	0·786
21	0·739	−0·311	−0·074	0·378	0·791
23	0·289	−0·200	0·409	0·083	0·298
27	0·510	0·232	0·375	−0·041	0·455
31	−0·065	0·555	−0·067	0·193	0·354
33	0·534	0·062	−0·600	−0·036	0·650
34	0·562	−0·180	−0·465	0·044	0·566

Webb's material by Garnett(5) and more fully described by Cattell(3). It is clear from the work of both these authors that this factor is strongly represented in traits 1 and 8, *cheerfulness* and *sense of humour*. As a preliminary measure we decided to pass one axis through the point representing trait 1, and at the same time to rotate the remaining axes in such a way as to preserve the orthogonality of the matrix. The first part of this task offers no difficulty; it is achieved by normalizing the co-ordinates of the point through which the axis is to be passed. This gives the direction cosines of the new axis referred to the axes of the centroid matrix. The first column of the new matrix is obtained by post-multiplication of M by

$$\begin{vmatrix} 0{\cdot}418 \\ 0{\cdot}808 \\ -0{\cdot}143 \\ -0{\cdot}389 \end{vmatrix}$$

The second part of the task is a little harder. The problem can be stated in geometric terms as follows. Given the r direction cosines of one axis in an r-dimensional space, it is required to find any orthogonal transformation which preserves the direction cosines of the given axis.

A general solution of this problem has been given by E. Schmidt(7), but as it involves the evaluation of r^2 determinants of the order $(r-1)$, we looked for a simpler method. Fortunately, *any* transformation will suffice which preserves the given axis and satisfies the criteria of orthogonality. In the four-dimensional case, with which we are immediately concerned, a solution can always be found by writing the given direction cosines a,

b, c, d, in four different orders, and adjusting the signs to satisfy the conditions of orthogonality, as in Table III. It will be seen that the cross

Table III

$$\begin{vmatrix} a & b & -c & -d \\ b & -a & d & -c \\ c & d & a & b \\ d & -c & -b & a \end{vmatrix}$$

product of any two columns in Table III is equal to zero, and if the sum of the squares of the elements in the first column is equal to unity, so also is the sum of the squares of every other column. This, of course, is not the only possible arrangement, but since we are concerned for the moment only with the first column, it is immaterial which of the many possible orthogonal transformations is employed.

The method we suggest here is of particular value when one wishes to make use of any of the current methods of locating a single axis, such as Thurstone's single hyperplane method, and at the same time to transform orthogonally. It may be of use therefore to show how it can be extended to any number of dimensions.

The solution for r dimensions depends on the fact that the product of two or more orthogonal transformations is itself an orthogonal transformation. If each of these involves the rotation of four axes by a transformation of the type shown in Table III, and if each of the r axes is rotated by at least one of these transformations, then each of the terms appearing in the several transformations is represented in the first column of the product. By equating the elements in this column to the given direction cosines, we arrive at a series of linear equations whose solution yields numerical values for all the terms used, and substitution of those values in the product gives the required orthogonal transformation. We present a solution of the seven-dimensional problem by way of illustration. Table IV shows that a seven-dimensional orthogonal transformation can be written in terms of eight values, and that each of these values is represented in the first column of the transformation product T. If the given direction cosines are a, b, c, d, e, f, g, put $x=a$, $y=b$, $z=c$, $pw=d$, $qw=e$, $rw=f$, $sw=g$. Then $w=\sqrt{\{(1-(a^2+b^2+c^2)\}}$, $p=d/w$, $q=e/w$, $r=f/w$, $s=g/w$. Substitution of the resulting values of x, y, z, w, p, q, r and s in T gives an orthogonal transformation in which the first column contains the given direction cosines.

The number of transformations that have to be multiplied together is $(r-1)/3$, or the nearest integral number above $(r-1)/3$, where r is the

Table IV

$$\begin{vmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & p & -q & -r & -s \\ 0 & 0 & 0 & q & p & s & -r \\ 0 & 0 & 0 & r & -s & p & q \\ 0 & 0 & 0 & s & r & -q & p \end{vmatrix} \cdot \begin{vmatrix} x & -y & -z & -w & 0 & 0 & 0 \\ y & x & w & -z & 0 & 0 & 0 \\ z & -w & x & y & 0 & 0 & 0 \\ w & z & -y & x & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{vmatrix}$$

T_1 T_2

$$= \begin{vmatrix} x & -y & -z & -w & 0 & 0 & 0 \\ y & x & w & -z & 0 & 0 & 0 \\ z & -w & x & y & 0 & 0 & 0 \\ pw & pz & -py & px & -q & -r & -s \\ qw & qz & -qy & qx & p & s & -r \\ rw & rz & -ry & rx & -s & p & q \\ sw & sz & -sy & sx & r & -q & p \end{vmatrix}$$

$T_1T_2=T$

number of dimensions. Thus when $r>7$, three or more transformations will be required. The first of these rotates the last four axes; in the second, the axes involved are the 4th to the 7th, counting from the rth backwards; and so on. It will be seen that the 4th last axis is rotated by both the first and second transformations, the 7th last by both the second and third, and so on. But if $(r-1)/3$ is not an integral number, there will be an overlap of 2 or 3 axes between the last and the second last of the transformations. If $r/3$ is an integral number, the work may be simplified by writing the last transformation in the form shown in Table V. Finally, if $(r+1)/3$ is an integral number, the last transformation may take the form shown in Table VI.

Table V

$$\begin{vmatrix} x & -y & -z & 0 & 0 & \dots & 0 \\ y & x & 0 & -z & 0 & \dots & 0 \\ z & 0 & x & y & 0 & \dots & 0 \\ 0 & z & -y & x & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & 1 & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 \end{vmatrix}$$

Table VI

$$\begin{vmatrix} x & -y & 0 & \dots & 0 \\ y & x & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & 1 & \dots \\ 0 & 0 & 0 & \dots & 1 \end{vmatrix}$$

The matrix M was multiplied by a transformation of the type shown in Table III. Having thus obtained an orthogonal matrix in which the first column represents the loads of the provisionally selected first factor, we plotted all the axes against one another in pairs. A number of rotations were tried out. Those which were retained are as follows:

(1) The second and fourth axes were rotated through a negative angle

of $\cos^{-1} 0{\cdot}939$. In this way the second axis was made to correspond closely to Webb's factor '*w*', traits 33 and 34 being given maximum weight at the positive end, and trait 2 maximum weight at the negative end.

(2) On considering the third axis along with the new fourth axis, we found that if trait 4 were given maximum weight at the negative end of the third axis, all the loadings on the fourth axis became positive or zero, which suggested that the fourth primary quality might be unipolar. To obtain this result we rotated the axes through a positive angle of 45°.

(3) The new second axis was taken together with the first axis, and rotated through a negative angle of $\cos^{-1} 0{\cdot}875$. This passes the second axis through the point standing for trait 2, giving it a maximum negative load. The justification for this procedure is that trait 2, *quick oscillation of mood*, is presumably relatively simple. It also has the merit of treating traits 33 and 34, *tendency not to abandon tasks*, etc., as more elementary than traits 20 and 21, *trustworthiness* and *conscientiousness*.

Table VII. *Transformation matrix* T

0·063	0·749	0·556	0·354
0·921	0·004	−0·282	0·268
0·186	−0·632	0·738	0·147
−0·336	−0·197	−0·259	0·884

The final transformation matrix, giving effect to all these steps, is presented in Table VII. Multiplying the centroid matrix *M* by the transformation matrix *T*, we get the factorial matrix *F*, shown in Table VIII.

Table VIII. *Factorial matrix* F

	I	II	III	IV	h^2
1	0·720	0·397	0·000	0·000	0·676
2	−0·007	−0·706	−0·025	0·067	0·504
3	−0·382	−0·569	0·066	−0·002	0·474
4	0·118	−0·465	−0·824	−0·110	0·921
5	0·257	0·401	0·502	0·215	0·525
6	0·042	−0·423	−0·646	0·157	0·623
8	0·792	0·178	0·016	0·108	0·671
11	0·294	−0·333	−0·571	−0·101	0·534
15	0·801	−0·179	−0·310	0·189	0·805
17	0·345	0·075	0·240	0·456	0·390
18	0·052	0·487	0·467	0·590	0·806
19	0·456	0·126	−0·037	0·659	0·659
20	−0·258	0·599	0·313	0·512	0·785
21	−0·381	0·525	0·346	0·501	0·792
23	−0·118	−0·059	0·497	0·182	0·298
27	0·329	0·154	0·505	0·261	0·455
31	0·430	−0·042	−0·292	0·287	0·354
33	−0·009	0·786	−0·158	0·086	0·650
34	−0·232	0·705	0·009	0·121	0·566

(4) *Conclusions as to qualities isolated*

The four primary qualities thus isolated may now be described. In view of the large amount of error necessarily incorporated in such a matrix, it seems unwise to treat as significant any factor loading of less than 0·2, or to use any trait with a load of much less than 0·4 for the purpose of describing and defining a primary quality.

At the positive end of the first axis, traits 15, 8 and 1, viz. *fondness for large social gatherings*, *degree of sense of humour*, and *general tendency to be cheerful*, have loads between 0·7 and 0·8; traits 19 and 31, *degree of corporate spirit*, and *bodily activity in pursuit of pleasure*, have loads of over 0·4; whereas at the negative end traits 3 and 21, *occasional extreme depression* and *conscientiousness*, have loads of 0·38. This factor is obviously akin to Cattell's Surgency-Desurgency, although it may not be absolutely identical with it. It can perhaps best be described by the antithesis between light-heartedness and serious-mindedness.

The second factor is heavily determined. On the positive side traits 33 and 34, *tendency not to abandon tasks*, etc., have loads of 0·7 or over. Traits 20 and 21, *trustworthiness*, and *conscientiousness*, have loads over 0·5, with trait 18, *tendency to do kindnesses on principle*, not far behind. Traits 5 and 1, *readiness to recover from anger*, and *tendency to be cheerful*, also have significant loads. On the negative side trait 2, *quick oscillation of mood*, has a load of over 0·7. Traits 3, *liability to extreme depression*, 4, *readiness to become angry*, and 6, *occasional extreme anger*, are all in excess of 0·4. *Eagerness for admiration*, with a load of 0·333, is probably also significant. This factor differs from Webb's '*w*' chiefly by giving more weight to steadiness, continuity and perseverance than to action from principle or purpose. Possibly the most appropriate name for it at present is Perseverance.

These first two factors are modifications of old friends. Our third and fourth factors are relatively new in psychological analysis.

The third factor is defined at the positive end by four traits with loads in the neighbourhood of 0·5. These are 27, *tact*, 5, *recovery from anger*, 23, *readiness to accept the sentiments of one's associates*, and 18, *kindness on principle*. *Conscientiousness* and *trustworthiness* lie between 0·3 and 0·4, and trait 17, *impulsive kindness*, with a load of 0·24, is also significant. At the negative end trait 4, *readiness to become angry*, has a load of over 0·8, and traits 6, *occasional extreme anger*, and 11, *eagerness for admiration*, are in the neighbourhood of 0·6. The loads of trait 15, *fondness for large social gatherings*, and trait 31, *bodily activity in pursuit of pleasure*, are

probably also significant. We have not met with an account of this quality in recent psychological literature, but something reasonably like it is described in I Corinthians xiii. 4–7. There it is called Charity, and we may let the name stand.

As already indicated, the fourth quality is unipolar. It is defined mainly by five traits, whose loads range from 0·659 to 0·456. These are 19, *degree of corporate spirit*, 18, *tendency to do kindnesses on principle*, 20, *trustworthiness*, 21, *conscientiousness*, and 17, *impulsive kindness*. This factor may be called Social Sensitiveness, which is not to be confused with Sociability, or anything resembling a herd instinct. This is clear from the fact that *fondness for large social gatherings* has a factor loading of less than 0·2, and that *eagerness for admiration* has an insignificant negative load.

Table IX shows the percentage of the variance of each trait attributable to the four common factors, and in the final column the percentage of the variance due to specific factors, including chance error. These factors are special only in the sense that they cannot be analysed by means of the data before us. If other traits are taken into account, further common factors may appear, and a fuller analysis will be possible. The need for further investigation is clearly indicated by the large values that u^2 has for some of the traits, notably 23, 31 and 17, where the unique variance is over 60%.

Table IX. *Percentage of variance attributable to common factors*

	I	II	III	IV	h^2	u^2
1	52	16	00	00	68	32
2	00	−50	00	00	50	50
3	−15	−32	00	00	47	53
4	01	−22	−68	−01	92	08
5	07	16	25	05	53	47
6	00	−18	−42	02	62	38
8	63	03	00	01	67	33
11	09	−11	−33	−01	53	47
15	64	−03	−10	04	81	19
17	12	00	06	21	39	61
18	00	24	22	35	81	19
19	21	02	00	43	66	34
20	−07	36	10	26	79	21
21	−15	28	12	25	79	21
23	−01	00	25	03	30	70
27	11	02	26	07	46	54
31	18	00	−09	08	35	65
33	00	62	−02	00	65	35
34	−05	50	00	01	57	43

The values of h^2 are copied from Table VIII, the occasional discrepancies between these values and the sum of the entries in columns I, II, III and IV above being due to the rounding off of numbers. The negative signs are included merely to show the direction in which the factors operate. They are, of course, neglected in calculating h^2.

In making this analysis we have been guided as much by psychological as by mathematical considerations. For example, it has been urged that it is advisable to aim at a solution that is mathematically unique, and that such a solution can be obtained by maximizing the number of zero entries in F. But a solution of this kind depends on the way in which the traits have been selected, and unless they are very numerous, and cover the whole field in an impartial fashion, no great importance can be attached

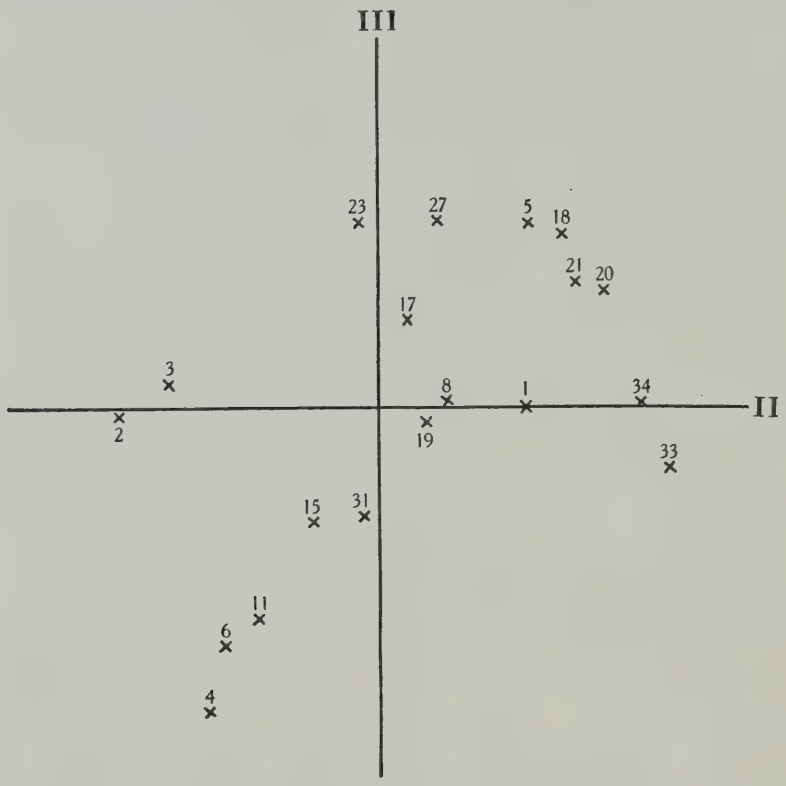


Fig. 2. Plot of columns II and III of F.

to their actual distribution in the common factor space. Of course considerations of the kind in question have not been altogether neglected, and it will be seen from Table IX that in each of the four columns there is a reasonable number of entries in the neighbourhood of zero. But on the other hand, in determining the position of the axes we have given more weight to psychological probability. A reference to Fig. 2 will illustrate the point. If the axes are rotated through a positive angle of 48°,

the loads of traits 4, 6 and 11 at the one end, and of traits 5, 18, 20 and 21 at the other end of the new second axis will be maximized, while their loads on the new third axis will be reduced to negligible quantities. This simplification, however, is gained at the expense of an increased complexity in traits 1, 2, 3, 33 and 34. We are thus faced with the choice between two analyses, which place the complexity with regard to these two factors in different groups of traits. We consider that traits 20 and 21 are more likely to be complex than 33 and 34, and that trait 2 has a greater claim to simplicity than trait 4. This has decided the position of the axes for us.

The analysis as it stands seems internally coherent, but there are gaps in it. For example, in Fig. 2 the second and fourth quadrants are virtually empty. But it is not difficult to suggest what qualities would help to fill them. Mr Pliable, lacking in '*w*' but probably endowed with some Charity, would almost certainly be found in the second quadrant; and over against him in the fourth quadrant we should expect to locate the obstinate, irritable, cantankerous fellow, along with his even more objectionable neighbour who vindictively nurses a grievance. It would appear, however, that these qualities did not flourish in the groups that Webb studied, and for analysis of them we must turn elsewhere.

REFERENCES

(1) Bernreuter, R. G. (1933). "The theory and construction of the Personality Inventory." *J. Soc. Psychol.* IV, 387.
(2) Burt, C. (1937). "Methods of factor-analysis with and without successive approximation." *Brit. J. Educ. Psychol.* VII, 172.
(3) Cattell, R. B. (1933). "Temperament tests: I. Temperament; II. Tests." *Brit. J. Psychol.* XXIII, 308; XXIV, 20.
—— (1936). "Temperament tests in clinical practice." *Brit. J. med. Psychol.* XVI, 43.
(4) Flanagan, J. C. (1935). *Factor Analysis in the Study of Personality.* Stanford University Press.
(5) Garnett, J. C. M. (1918). "General ability, cleverness and purpose." *Brit. J. Psychol.* IX, 345.
(6) Kelley, T. L. (1935). *Essential Traits of Mental Life.* Harvard University Press.
(7) Schmidt, E. (1908). "Über die Auflösung linearer Gleichungen mit unendlich vielen Unbekannten." *Rendiconti del Circolo Matematico di Palmero,* XXV, 53.
(8) Spearman, C. (1927). *The Abilities of Man.* Macmillan.
(9) Thurstone, L. L. (1935). *The Vectors of Mind.* University of Chicago Press.
(10) —— (1938). *Primary Mental Abilities.* University of Chicago Press.
(11) Webb, E. (1915). "Character and intelligence." *Brit. J. Psychol. Monogr. Suppl.* no. 3.

(*Manuscript received* 16 *November* 1938)

TWELFTH INTERNATIONAL CONGRESS OF PSYCHOLOGY

Preliminary Notice

The Twelfth International Congress of Psychology will be held in Edinburgh, Scotland, from 22 to 27 July 1940. The Honorary President of the Congress will be the Rt Hon. John Colville, M.P., Secretary of State for Scotland.

The Congress fee has been fixed at thirty shillings (£1. 10*s.*) sterling for active members and fifteen shillings (15*s.*) sterling for associates.

Arrangements are being made by which a large proportion of the members can be accommodated in University hostels. The inclusive charge for such accommodation will be about fifty shillings (£2. 10*s.*), excluding midday lunches, arrangements for which will be made elsewhere. Members who do not desire accommodation in the hostels can easily secure accommodation in hotels and private hotels, of which Edinburgh has a large number.

When the Committee of Organization is fully constituted, formal invitations will be issued. It will, however, greatly facilitate arrangements if as many as possible will let the General Secretary know now that their attendance at the Congress is probable. In the meantime subjects for symposia, general discussion and lectures are under consideration by a Preliminary Arrangements Committee.

A volume of Proceedings of the Congress, containing abstracts of papers read, will be published, the cost of which is included in the Congress fee for active members.

James Drever, *President of Committee of Organization.*

Godfrey Thomson, *General Secretary.*

(Moray House, Edinburgh 8)

PUBLICATIONS RECENTLY RECEIVED

The Factorial Analysis of Human Ability. By G. H. Thomson. University of London Press. 1939. Pp. xv + 326. 16*s.* net.

Since the pioneer work of Professor Charles Spearman in 1904 a tremendous amount of work has been carried out on the factorial analysis of psychological test scores and of all that has been written on this subject a great deal has been controversial. One of the foremost critics of the Spearman school of analysis was Professor Godfrey H. Thomson, the author of the book under review, and although the points at issue cannot yet be regarded as settled, the statement of the present position made in Professor Thomson's book will be of great value both to those who are attracted by the mathematical approach to psychology and also to those who regard it with suspicion.

Factorial analysis may be considered from three points of view: first, there is the question of the mathematical validity of the various methods of analysis in present use; secondly, there is the difficulty of the actual mathematical calculations involved; and thirdly, and of paramount importance to the psychologist, there is the question as to whether factors isolated mathematically have any psychological existence or can indeed be of any use to the psychologist either theoretically or practically.

As regards the theoretical nature and mathematical validity of the various methods of factorial analysis, only an advanced mathematician is capable of appreciating and judging and the non-mathematical psychologist must perforce accept the ruling of the mathematicians on this point. Professor Thomson deals with this question in a manner calculated to satisfy the advanced mathematician and to instruct the person of average mathematical ability. To the psychologist the interesting fact is made quite clear that there is as yet no unique method of factorizing a set of test scores but that in each of the methods at present in use resort is made to some assumption enabling factors to be estimated or to some device such as "purifying" the battery of tests by rejecting tests which do not yield correlation coefficients which fit into some hierarchical matrix. The psychologist who intends to use some particular method of factorial analysis should therefore ascertain precisely what assumptions he is making therein; he will find these clearly stated in this book.

Turning to the matter of the actual calculations involved in factorial analysis we are greeted in the preface of the book with the statement that "this book has been written so that it can, it is hoped, be read by those who have no mathematics beyond the usual secondary school knowledge." Now it is true that the examples of calculations given in the book involve only elementary arithmetic and algebra, but at the same time the reader must be capable of sustained mathematical thinking in order to make any headway at all with the subject. The mathematical methods used are not such as are normally included in a secondary school curriculum and the reader must be completely familiar with the meaning of such terms as "variance" and "correlation". Moreover, the author himself says that the serious student should have some knowledge of matrix algebra and that is a branch of mathematics which is rarely dealt with even in a university mathematics course. One may conclude, therefore, that whilst the person of average mathematical ability might slavishly follow the arithmetical examples provided and even copy them using his own data, yet only one capable of clear and sustained mathematical thought would be able to understand what he was doing at all stages of the analysis. Such a one will, however, find in this book adequate and clear explanations of the analytic procedures adopted by Spearman, Thurstone, Hotelling, Burt and others, with actual worked examples and accounts of the various "short-cut" methods which have been devised.

The third aspect of factorial analysis, the question of the psychological meaning of the factors emerging from mathematical manipulation, is one which has the chief interest for the psychologist. Professor Thomson has a good deal to say on this point. Perhaps the most significant statement he makes may be quoted in full. He writes

(p. 267): "In brief, then, the author's attitude is that he does not believe in factors if any degree of real existence is attributed to them; but that, of course, he recognizes that any set of correlated human abilities can always be described mathematically by a number of variables or "factors", *and that in many ways*, among which no doubt some will be more useful or more elegant or more sparing of unnecessary hypotheses. But the mind is very much more complex, and also very much more of an integrated whole, than any naïve interpretation of any one mathematical analysis might lead a reader to suppose. Far from being divided up into "unitary factors", the mind is a rich, comparatively undifferentiated complex of innumerable influences—on the physiological side an intricate network of possibilities of intercommunications. Factors are fluid descriptive mathematical coefficients, changing both with the tests used and with the samples of persons, unless we take refuge in sheer definition based upon psychological judgment, which definition would have to specify the particular battery of tests, and the sample of persons, as well as the method of analysis, in order to fix any factor." Coming from the pen of Professor Thomson, the passage quoted is extremely significant and should make the psychologist very cautious in accepting and interpreting any mathematical factors which he may isolate. There is a grave danger, as is pointed out in this book, of reifying mathematical factors—a danger which, though usually avoided by the leading exponents of factorial analysis, is frequently positively courted by their enthusiastic disciples. Of what use then is factorial analysis to the psychologist? the reader will ask. It is rather difficult to find an answer from this book except that the author is sceptical of the utility and pessimistic of the future of factor analysis. He concludes, however, that the factor theory will still continue in spite of all its dangers and imperfections, partly on account of its intrinsic interest and partly because it has undeniably been a guide and spur to many investigators.

By examining and explaining *in extenso* the dangers and imperfections inherent in factorial analysis Professor Thomson has rendered a great service to the psychological investigator who has neither the time to study the vast body of original documents on the subject nor the mathematical ability to criticise and appraise the methods for himself. But the book is not easy reading for the non-mathematician.

E. G. C.

American Psychology before William James. By Jay Wharton Fay. Rutgers University Press. 1939. Pp. x+240. $3.00.

As its title indicates, this book is of historical interest, and, on the whole, of minor historical interest only. Miss Fay has no difficulty in showing that both American writers and others have neglected American psychology before William James, but it seems likely that the perusal of her interesting and brightly written pages will leave most readers still of the same opinion that very little of original and striking character was achieved in America during the long period with which her book is concerned. She divides the whole period into three: the period of theology and moral philosophy (1640–76); that of intellectual philosophy (1776–1827), and that of British and German influences (1861–90). The period from 1827–61 is hailed as "the era of American textbooks": 1861 seems rather an early terminal point to set for this. On the whole she has done her work well, and must have read a vast quantity of literature, most of it dull. But the impression remains that until William James brought his genius into the field of psychology, the American contributions were derivative and on the whole more pretentious than helpful.

English, and especially Cambridge readers (possibly provided they do not belong to the college mentioned) will be delighted with an early paragraph: "The Puritans made prompt preparations for elementary education, and the need for preparation of ministers of the Gospel led to the founding of the first college, modelled after Magdalene at Cambridge, a Puritan institution in which many of the New England colonists, including John Harvard, had studied." But this is not a fair sample of the author's accuracy.

Logical Aspects of Educational Measurement. By B. Othanel Smith. New York: Columbia University Press. In England, Oxford University Press: Humphrey Milford, 1938. Pp. x+182. 12*s.* 6*d.*

The main thesis of this book is to define what is meant by measurement, to enunciate the essential principles on which the construction of measuring instruments must be based, and to examine critically the question of whether learning has the property of being measurable and, if so, whether the instruments that have been used for its measurement are valid. The author states that measurement is the quantitative evaluation of a property by means of an instrument which is constructed in accordance with certain general principles, and that to be completely measurable a property must be capable of being described as more or less in amount, it must be such that a property of like kind can be found equal to it, and it must be capable of addition. The author denies that equal units have been established in measuring educational attainment and shows that the many measuring instruments which have been used in educational and psychological measurement do not satisfy the axioms of addition. It follows, therefore, that learning and educational attainment cannot be completely measured but only partially measured in a manner similar to the measurement of temperature by means of a thermometer.

The author forecasts that the future of educational measurement will be in two directions: "First, it will seek for those independent and homogeneous elements with which it can deal most successfully; and second, it will rest content with the degree of accuracy it has attained and will attempt to adapt its techniques and procedures to the requirements of qualitative evaluation." The problem discussed in this book is one which is of importance to the increasing number of psychologists who rely on test scores and statistics for the basis of their theories. The exposition is reasonably clear but the reader is irritated by innumerable footnotes, most of which could be avoided by the usual system of reference to the comprehensive bibliography at the end of the book.

Psychophysiologie der Tastsinne. By Emil v. Skramlik. Leipzig: Akademische Verlagsanstalt. 1937. Pp. xiii+935, 308 illustrations. R.M. 78.

The book attempts to give a comprehensive survey of the facts and theories of the tactual senses, the discussion being centred about the senses of pressure, temperature, pain and kinaesthesis, and the cooperation of these four senses with one another. A particularly detailed treatment is accorded to those fields in which experimental contributions have been made by the author himself and by his pupils. Although the amount of literature covered is tremendous, the work cannot be considered as completely exhaustive, as I was able to determine, for instance, when I checked the discussion of the problems of vibratory sensivity. Wherever specific psychological problems are touched upon, one feels that the treatment does not do justice to the "totality" point of view which nowadays receives, and deserves, so much emphasis. Nevertheless, the psychologist owes the author a debt of gratitude for having undertaken the gigantic task of presenting the facts of tactual sensitivity in comprehensive form. Every future worker in this field will find it advisable to go to Skramlik for purposes of orientation.

David Katz.

Abnormal Speech. By E. J. Boome, H. M. S. Baines and D. G. Harries. London: Methuen. 1939. Pp. xiv+162. 6*s.* net.

This book should be of considerable use, especially to the practising physician, provided it is read with care and is not regarded as a complete and final treatment of the subject. It contains a very useful classification of disorders of speech, both

organic and psychogenic, although some of it, particularly the classification of aphasic and allied defects, seems somewhat out of date. Further, it is distinctly practical in bent and indicates plenty of exercises, most of them excellent and clearly described, for improving utterance. On the whole the least satisfactory section concerns stammering and stuttering, because it is suggested that the basic successful treatment is simply by relaxation. The authors do, indeed, take care to say that this treatment is essentially an early device, and that the defects may, if they are at all long standing, be complicated by many psychological difficulties which render treatment hard and recovery far more prolonged. But this is not emphasized sufficiently, and there are plenty of cases, even in the early stages, where the relaxation recommended will be found very hard to obtain unless special measures are taken. Still, if it is read carefully, and checked against actual experience, the book should do a lot of good. It contains much information concisely arranged, and some interesting statistics, as well as an abundance of suggested practical exercises.

Mental Tests. By FRANK N. FREEMAN. London: G. Harrap. 1939. Pp. xi+460. 10*s.* 6*d.* net.

Professor Freeman's text-book is already well-known and has been found widely useful. It has been thoroughly revised and brought up to date. It still remains in the main a factual book, concerned especially to describe and to illustrate a wide variety of tests, to give sufficient information about common methods of marking and scoring, and very briefly to characterize the main theories in the field. It is an extremely efficient piece of work, and should be of wide use in training college and university courses.

The Dream World. By R. L. MÉGROZ. London: The Bodley Head. 1939. Pp. xiv+319. 10*s.* 6*d.* net.

Here is a most excellent book in which to dip and read. It could almost be called Dreams with Notes, for although it has an introduction and a historical commentary on theories, it is mainly full of dream descriptions and these are gathered both from famous and from unknown people. Very queer and interesting many of these dreams are, and the psychologist who pleases can read anywhere and interpret as he likes, or if he likes better he need not interpret at all, but can simply enjoy. As a matter of fact the author writes with learning and with skill and his book is much more than a mere compilation. The book is all the better because, though it is clear that the author knows a lot about his subject, he has most skilfully avoided all elaborate and technical treatment.

Conscious Orientation. By J. H. VAN DER HOOP. London: Kegan Paul, Trubner, Trench and Co., Ltd. 1939. Pp. x+352. 15*s.* net.

There is little doubt of the difficulty in arriving at some adequate method of describing and classifying qualitative differences in personality; and the concept of types seems in many ways to offer a useful way of surmounting this difficulty. Jung's dual, extravert-introvert, classification has always been one of the most popular of these typologies; but his further classification into sensation, feeling, intuition and thinking types is much harder to grasp, and it has been adopted by few other psychologists. Dr Van der Hoop, however, finds that with suitable modification it may usefully be employed in the understanding and description of different types of personality as the latter appears in its conscious orientation towards the environment, both in the normal individual, and also in the psychotic and neurotic. But he substitutes for the sensation type, the instinctive type—comprising those who react immediately not

merely in response to their sensations, but also in accordance with their instinctual and emotional needs. The feeling type is guided by more consistent and enduring affective attitudes, in the nature of McDougall's sentiments, in which a variety of emotional reactions are combined. Characteristic differences are observed according as the orientation of the personality is introverted or extraverted.

Dr Van der Hoop then demonstrates how such an eight-fold classification may be used to gain an understanding of the structure of the personality which is particularly useful to the psycho-pathologist. In the neurotic, this structure conforms to those of the normal type, and determines to a great extent the nature of the abnormal behaviour set up by some excessive environmental strain of the kind dealt with by Freud. A better understanding of the psychotic is obtained, however, if some abnormal deviation is recognized from the structural type.

It seems probable that this typological scheme may assist both the psychologist and the psycho-pathologist in the understanding of those qualitative differences in the structure of personality which are so baffling and so obstructive to the assessment and description of individual personality. As to the ultimate reality and the philosophical validity of his scheme, we must for the present remain doubtful.

Mis-Mated: The Psychology of the Unhappy Marriage. By JOHN F. PETERS. London: John Bale Sons and Curnow Ltd. 1938. Pp. 213. 7*s.* 6*d.* net.

"Are you aware", ask the publishers, "that your position in your family, in relation to your brothers and sisters, determines your chances of making a happy marriage? Are you aware that if (for example) you are an only son without a younger sister, you will be unhappy if you marry a woman with elder brothers, etc.?" This perhaps sufficiently indicates the topic of the book, which is written with immense high spirits, tremendous confidence, and much illustration, by a man who many times points out that he is an only son and hence an introvert, and who obviously has a great interest in other people. No brief summary is possible. The whole rests upon the complete acceptance of the introvert-extravert classification of human beings—together with whatever border types are convenient. Introversion or extraversion are asserted to be determined early by position in the family, and whether a person is happy in marriage or not depends on his or her finding a mate who has the right relation of introversion or extraversion to himself or herself. Hence the great thing to do is not to fall in love until one has discovered how many brother and sisters the potential loved one had, and how they were spaced out. There is a lot of ingenious illustration drawn from a study of the married happiness or unhappiness of famous people, and from more or less analogous evidence.